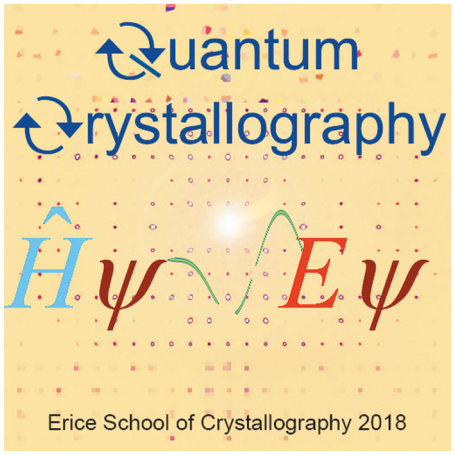
International School of Crystallography

Director: Sir Tom Blundell, FRS FMedSci



52nd Course: Quantum Crystallography

Erice, Italy • 1-10 June 2018

Course Directors: Dylan Jayatilaka · Piero Macchi

Programme, Lecture Notes & Poster Abstracts

Ettore Majorana Foundation and Centre for Scientific Culture
President and Director: Professor Antonino Zichichi

Acknowledgements

This Course is supported by

Organisation for the Prohibition of Chemical Weapons (OPCW)
European Crystallographic Association
Associazione Italiana di Cristallografia
Dectris
International Union of Crystallography
Crystals (Open Access Journal of Crystallography)
The Cambridge Crystallographic Data Centre (CCDC)
Malvern Panalytical
MSD
Universität Bern
CRYSTAL
BRUKER
STOE

Organisers

Course Directors

Dylan Jayatilaka

Piero Macchi

Local Organisers

Paola Spadon

Annalisa Guerri

IT Support

Fred Boyle

Erin Davis

Organisers

Orange scarves

Andrea Giaccherini

Gearóid Mangan

Valentina Marcheselli

Paolo Pio Mazzeo

Fabio Montisci

Fabio Nicoli

Giovanna Scapin

Martin Schmidt

Programme

Saturday, 2 June 2018

8:30-9:00 **Introduction to both courses**
A. Guerri, P. Spadon, E. Davis, J. Hadermann, L. Palatinus, A. Stewart, D. Jayatilaka, P. Macchi

9:00-9:45 **Joke Hadermann**
Fundamental Crystallography (essential for TEM users)

9:45-10:30 **Lou Massa**
Quantum Crystallography, Electron Density, and the Kernel Energy Method

10:30-11:00 ***Coffee break***

11:00-12:30 **Jan Pieter Abrahams**
General theory of diffraction

12:30-14:30 ***Lunch***

14:30-15:15 **Jacob Overgaard**
Accurate X-ray diffraction Measurements

15:15-16:00 **Anders Madsen**
Dynamics in crystals in the context of quantum crystallography

16:00-16:30 ***Coffee break***

16:30-17:15 **Piero Macchi**
Atom centered multipolar expansion of the charge density

17:15-18:00 **Simon Grabowsky**
Hirshfeld Atom Refinement

18:00-18:30 **Introduction to Erice** (Martin Schmidt)

20:00 ***Sicilian Dinner***

Sunday, 3 June 2018

9:00-10:30 **Concomitant Workshops:**
Data treatment for single crystal X-ray diffraction (multiple programs demonstrated); Multipolar refinement of charge density with MOPRO; Multipolar refinement of charge density with XD2016 and MoleCoolQT; Distributed Atomic Polarizabilities with PolaBer; Hirshfeld Atom Refinement (multiple programs demonstrated)

10:30-11:00 ***Coffee break***

11:00-12:30 **Concomitant Workshops:**
Data treatment for single crystal X-ray diffraction (multiple programs demonstrated); Multipolar refinement of charge density with MOPRO; Multipolar refinement of charge density with XD2016 and MoleCoolQT; Distributed Atomic Polarizabilities with PolaBer; Hirshfeld Atom Refinement (multiple programs demonstrated)

12:30-14:30 ***Lunch***

14:30-15:15 **Nicolas Claiser**
Neutron diffraction and spin density multipolar model

15:15-16:00 **Maxime Deutsch**
Neutron diffraction and spin density multipolar model

16:00-16:30 ***Coffee break***

16:30-17:15 **Martin Rahm**
Experimental Quantum Chemistry

17:15-18:00 **Dylan Jayatilaka**
X-ray constrained wave functions part 1

18:00-20:00 **Poster Session 1 (odd numbers)**

Monday, 4 June 2018

9:00-9:45 **Peter Müller**
Crystal Structure Refinement I

9:45-10:30 **Philip Nakashima**
QCBED Lecture 1: QCBED – A Nexus Between Quantum and Electron Crystallography

10:30-11:00 ***Coffee break***

11:00-12:30 **Poster Presentations • Rising stars in Quantum Crystallography:**

11:10-11:20 **Lorraine Andrade Malaspina**
HAR-ELMO - Fast and accurate Hirshfeld Atom Refinement

11:20-11:30 **Emil Damgaard-Møller**
A charge density study of a linear dialkyl Co(II) complex showing an unprecedented non-Aufbau electronic ground state

11:30-11:40 **Max Davidson**
General-unrestricted extremely localised molecular orbital (gELMO) wavefunctions for Hirshfeld atom refinement (HAR)

11:40-11:50 **Jonathan Du**
Wavefunctions fitted to charge density data – exploring the effects of data quality and obtainable properties

11:50-12:00 **Michelle Ernst**
Bonding in polyiodides

12:00-12:10 **Anna Krawczuk**
Influence of chosen synthons on the polarizabilities of functional groups

12:10-12:20 **Alexey Kuzmin**
Experimental observation of Jahn-Teller distortions in π-conjugated high symmetry systems: C_{60}^{n-} and $[MPc]^{n-}$ anions

12:30 **EXCURSION 1**

Tuesday, 5 June 2018

9:00-10:30 **Concomitant Workshops:**
Data treatment for single crystal X-ray diffraction (multiple programs demonstrated); Multipolar refinement of charge density with MOPRO; Multipolar refinement of charge density with XD2016 and MoleCoolQT; Wavefunction refinement with TONTO

10:30-11:00 ***Coffee break***

11:00-12:30 **Concomitant Workshops:**
Data treatment for single crystal X-ray diffraction (multiple programs demonstrated); Multipolar refinement of charge density with MOPRO; Multipolar refinement of charge density with XD2016 and MoleCoolQT; Wavefunction refinement with TONTO

12:30-14:30 ***Lunch***

14:30-15:15 **Bartolomeo Civalleri**
Periodic systems: models and strategies

15:15-16:00 **Paolo Giannozzi**
Introduction to density-functional theory and the plane-wave pseudopotential method

16:00-16:30 ***Coffee break***

16:30-17:15 **Alessandro Erba**
From Energy and Wavefunction to Advanced Properties of Solids

17:15-18:00 **Marcus Neumann**
Dispersion corrected DFT methods and crystal structure predictions

18:00-18:30 **Poster Presentations • Rising stars in Quantum Crystallography**

18:00-18:10 **Francesca Peccati**
Overcoming distrust in solid state simulations: the case of cell parameters

18:10-18:20 **Stefano Racioppi**
Electronic properties and bonding in azolate based coordination polymers

18:20-18:30 **Lucy Saunders**
Experimental charge density studies of short strong hydrogen bonds (SSHBs) with potential proton migration behaviour on I19, Diamond Light Source

Wednesday, 6 June 2018

9:00-10:30 **Concomitant Workshops:**
Multipolar refinement of charge and spin density with MOLLY-X-N; Multipolar refinement of charge density with XD2016 and MoleCoolQT; Crystal Orbital Calculations with CRYSTAL2017; Periodic DFT calculations with Quantum Expresso

10:30-11:00 ***Coffee break***

11:00-12:30 **Concomitant Workshops:**
Multipolar refinement of charge and spin density with MOLLY-X-N; Multipolar refinement of charge density with XD2016 and MoleCoolQT; Crystal Orbital Calculations with CRYSTAL2017; Periodic DFT calculations with Quantum Expresso

12:30-14:30 ***Lunch***

14:30-15:15 **Alessandro Genoni**
Extremely localized molecular orbitals in quantum crystalography

15:15-16:00 **Piero Macchi**
Atomic polarizabilities and dielectric properties

16:00-16:30 ***Coffee break***

16:30-17:15 **Julia Contreras**
Quantum Topology

17:15-18:00 **Poster Presentations • Rising stars in Quantum Crystallography**

17:20-17:30 **Rebecca Scatena**
Electron density and dielectric properties of highly porous MOFs

17:30-17:40 **Bjarke Svane**
Electron densities of organic molecular crystals from powder X-ray diffraction

17:40-17:50 **Daniel Tchoń**
Application of HAR to incomplete, high pressure data

18:00-20:00 **Poster Session 2 (even numbers)**

Thursday, 7 June 2018

9:00-10:30 **Concomitant Workshops:**
Multipolar refinement of charge density with MOPRO; Crystal Orbital Calculations with CRYSTAL2017; Distributed Atomic Polarizabilities with PolaBer; Wavefunction refinement with TONTO

10:30-11:00 ***Coffee break***

11:00-12:30 **Concomitant Workshops:**
Multipolar refinement of charge density with MOPRO; Crystal Orbital Calculations with CRYSTAL2017; Distributed Atomic Polarizabilities with PolaBer; Wavefunction refinement with TONTO

12:30 **EXCURSION 2**

Friday, 8 June 2018

9:00-10:30 **Concomitant Workshops:**
Periodic DFT calculations with Quantum Expresso; Hirshfeld Atom Refinement (multiple programs demonstrated); X-ray constrained extremely localized molecular orbitals

10:30-11:00 ***Coffee break***

11:00-12:30 **Concomitant Workshops:**
Periodic DFT calculations with Quantum Expresso; Hirshfeld Atom Refinement (multiple programs demonstrated); X-ray constrained extremely localized molecular orbitals

12:30-14:30 ***Lunch***

14:30-15:15 **Ulf Ryde**
Quantum refinement for biological applications

15:15-16:00 **Kenneth Merz**
Quantum refinement for biological applications

16:00-16:30 ***Coffee break***

16:30-17:15 **Benoit Guillot**
Experimental charge density studies in biomolecules

17:15-18:00 **Birger Dittrich**
Transferable Electron Densities

18:00-18:30 **Poster Presentations • Rising stars in Quantum Crystallography**

18:00-18:10 **Alena Vishina**
Electronegativity in band structure calculations

18:10-18:20 **Marcel Vöst**
On the compressibility of C-H bonds in late transition metal alkyls

18:20-18:30 **Erna Wieduwilt**
Heavy meets light - A systematic study of the bond between hydrogen and heavy elements

Saturday, 9 June 2018

9:00-9:45	**Dylan Jayatilaka** X-ray constrained wave functions part 2
9:45-10:30	**Wolfgang Scherer** Organometallic bonding (concepts) under pressure
10:30-11:00	***Coffee break***
11:00-11:45	**Franz Giessibl** Atomic force microscopy – microscopy with ultimate resolution
11:45-12:30	**Thomas Elsässer** Femtosecond X-ray diffraction
12:30-14:30	***Lunch***
14:30-15:15	**Mark Spackman** Quantum Crystallography and Crystal Engineering: Experimental lattice energies from X-ray diffraction data?
15:15-16:00	**Round Table, The Future of Quantum Crystallography introduction by Bo Iversen**
16:00-16:30	***Coffee break***
16:30-17:15	**Round Table, The Future of Quantum Crystallography (few short talks)**
17:15-18:00	**Joint Closing Remarks**
20:00	***Farewell Party***

Invited Speakers' Contributions

Fundamental Crystallography (essential for TEM users)

Joke A.M. Hadermann
University of Antwerp, Antwerp, Belgium
joke.hadermann@uantwerpen.be

Doing TEM, you will encounter different types of patterns and images. It is impossible to interpret these wisely without knowing crystallography. Why are the reflections in exactly those positions? Why is one reflection brighter than another? How does the symmetry in your electron diffraction pattern or high resolution image correlate with the symmetry in your structure?

1. Construction of the direct and reciprocal lattice

How images are formed and the details of diffraction you will see in other lectures in the course (if you are in the Electron crystallography course). For this lecture, it is sufficient to explain the geometric construction of the reciprocal lattice, to indicate its relation to the diffraction patterns of a structure. What you see on electron diffraction patterns corresponds to sections through the reciprocal lattice of the studied material. The reflections can be sharp as in selected area electron diffraction patterns or discs with contrast variations inside as in convergent beam electron diffraction. You might also be working with Fourier transforms of images. To understand each of these patterns, you need to be aware of how the reciprocal lattice is constructed.

A structure is made up of a lattice repeating itself in direct space. The lattice is decorated with atoms, forming a crystal structure. Each lattice point is equivalent (thus has the same atom decoration) and parallel directions have equal distances between lattice points. Basis vectors $\vec{a}, \vec{b}, \vec{c}$ can be chosen in different ways, but all lattice points need to be able to be described as a linear combination of these three basis vectors (Figure 1). Note that for aperiodic structures such as incommensurately modulated compounds or quasicrystals you need more dimensions than three. In this short refresher, we will not go into those advanced cases.

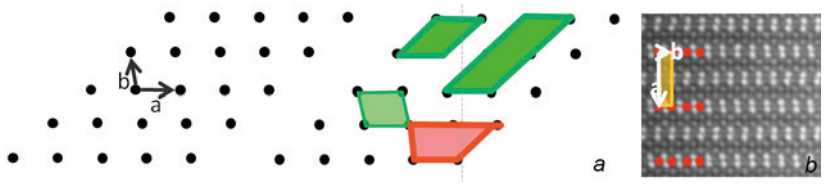


Figure 1. a) Example of a lattice, with the in-plane basis vectors a and b indicated (left) and valid alternative unit cells in green, and invalid example unit cell in red (right). b) A high resolution HAADF-STEM image on which similarly a valid unit cell and the corresponding basis vectors have been indicated.

A multitude of planes can be indicated as passing in different ways through the atoms. These planes are labelled with an index (hkl) obtained by marking where the plane cuts the different axes (Figure 2a), i.e. at points m, n and p for resp. the a, b and c-axis, reversing them to 1/m 1/n 1/p and taking the smallest set of integers with the same ratios. For example, if the plane cuts the a-axis at x=2, the b-axis at y=1 and the c-axis at c=3, the reverses are 1/2 1/1 1/3 and the index is (362). A single set of parallel planes is indicated by round brackets, (hkl). If you see accolades {hkl} it means (hkl) and all planes equivalent with (hkl).
For each set of planes (hkl) in a structure, you construct the reciprocal lattice point hkl by drawing the perpendicular to this set of planes (hkl), and positioning the point at a distance from the central beam that is the reverse of the distance d between the planes (multiplied by a consistent scale factor for drawing) (Figure 2b). The index of the corresponding reciprocal lattice point will be hkl, same number, no brackets. For example, the plane drawn in Figure 2a will correspond with a reciprocal lattice point 362.

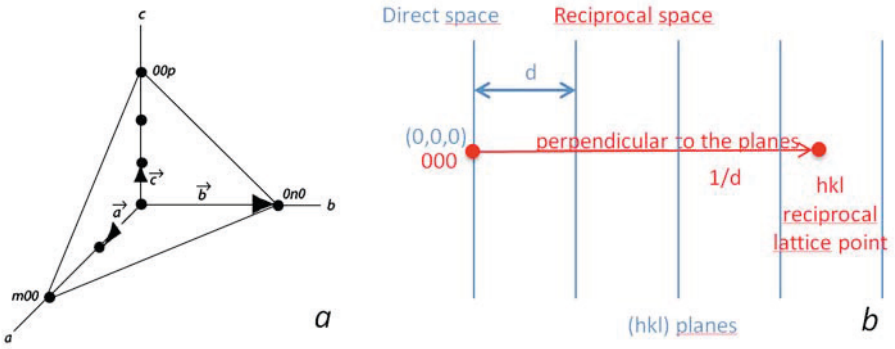


Figure 2. a) Planes are indexed using their intersections with the axes. b) Construction of reciprocal lattice points. Direct space elements are given in blue, reciprocal space elements in red.

As you have a threedimensional set of planes, you will also get a threedimensional set of points, forming a lattice. Also for the reciprocal lattice, parallel directions have equal distances between points. At this point in our consideration, all lattice points are still equivalent (an "unweighted lattice"), having the same intensity, we just take into account the geometry of the structure and the reciprocal lattice.
Any diffraction technique samples this reciprocal lattice, electron diffraction for example takes sections through this reciprocal lattice (Figure 3), and by putting all sections together you can then reconstruct the whole reciprocal lattice, with correct distances between the lattice points. The sections which we see on electron diffraction patterns are "zones". A zone is a collection of planes in direct space that have a common direction. The common direction is called the zone axis [uvw] ([ ] indicate a direction, < > indicate several equivalent directions). A zone with zone axis [uvw] in direct space will result in a set of reflections all lying in the same plane in reciprocal space which is therefore indicated as the zone [uvw]. In conventional electron diffraction

experiments you can only obtain zones that pass through the central beam (000 of the reciprocal lattice), in electron diffraction tomography reconstructions you can choose to see any reconstructed zone.

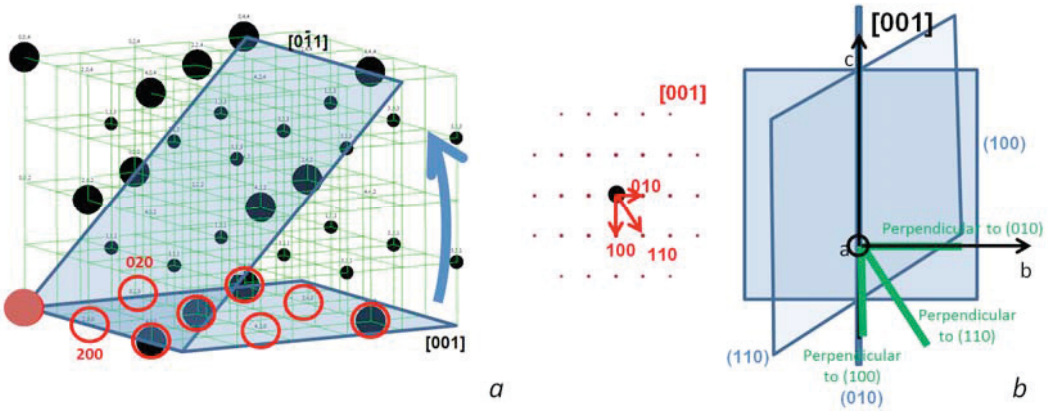


Figure 3. a) Threedimensional lattice of reciprocal lattice points (example is CaF_2, space group Fm-3m).The blue planes indicate two examples of sections through the lattice, corresponding to the zones [001] and [0-11]. b) Scheme of the relation between a [zone] electron diffraction pattern and the definition of a zone as the collection of planes with a common direction. Example shows zone [001], which includes (not all planes drawn) a.o. the planes (100), (110) and (001). The corresponding electron diffraction pattern will thus a.o. contain the reflections 100, 110 and 010.

The basis vectors a*,b*, c* of the reciprocal lattice are vectors with which you can describe all lattice points with a linear combination hkl=ha*+kb*+lc*. Planes with a certain index (hkl) will always have the same interplanar spacing for the chosen basis. Thus by indexing diffraction patterns you can derive the reciprocal cell parameters. It is again possible to choose different bases for describing the lattice. Once you have found a basis (a*, b*, c*, α*, β*, γ*) with which you can assign to each reflection an index hkl, with h, k and l all integers, you can calculate from these reciprocal basis vectors the cell parameters (a, b, c, α, β, γ) of your unit cell in direct space, using the general relations (with V the volume of the unit cell)

$$\overrightarrow{a*} = \frac{\vec{b}\times\vec{c}}{(\vec{a}\times\vec{b}).\vec{c}} \qquad \text{and} \qquad |\overrightarrow{a*}| = a* = \frac{1}{d_{100}} = \frac{bc\sin\alpha}{V}$$

(cyclic permutations give similar for b and c). The equations between the three aspects (d_{hkl}, (hkl), cell parameters a-b-c and angles between the axes $\alpha - \beta - \gamma$) that can be found in most books on crystallography or TEM.

To each choice of basis for the structure in direct space corresponds one basis choice in reciprocal space and vice versa. The positions of the reflections in the electron diffraction patterns are define a unique lattice. For this unique lattice, different descriptions (bases) can be chosen, but the lattice itself does not change because of different choices for the description. Transformations between chosen bases can be made using simple matrix relations, with a

transformation matrix P. This P or its inverse can then also be used to transform coordinates, indices of reflections, etc. More details can be found in the International Tables of Crystallography, volume A, section 5.

2. Intensities of reflections

The previous section explains why reflections in diffraction patterns or Fourier transforms are at specific positions, and how indexing (assigning a consistent set of hkl to the different reflections) can give you the cell parameters of your structure. While the positions of the reflections depend on the positions of the different planes relative to each other, the intensities depend on what is the occupation of the different planes. This will be explained in more detail in the lectures on diffraction in general, and can for example be expressed using the structure factor. The terms of the structure factor depend on the scattering factor of the atoms at certain positions, multiplied by a factor indicating the position of the atoms (x_j, y_j, z_j) and the reflection (hkl) being considered.

$$FH = F(hkl) = \sum_{j=1}^{N} f_j e^{2\pi i \vec{\rho_j} \vec{B_H}} = \sum_{j=1}^{N} f_j e^{2\pi i (hx_j + ky_j + lxz_j)}$$

Several reflections can have equivalent intensities, and also several sets of reflections can have zero intensities. This all depends on the symmetry of the structure. If several terms in this equation of the structure factor have the same amplitude, but opposite phase, they will cancel each other out, as in destructive interference of waves. If they do not have the same amplitude, they will simply strengthen or decrease the overall intensity of the reflection, depending on the phase difference between the terms. This relation is the same for the different diffraction techniques, but the scattering factors will differ, thus you will get different theoretical intensities for X-ray diffraction than for electron diffraction.

By analysing which extinctions are systematically absent, you can get information on the translation symmetry of the structure of your compound. By analysing the symmetry among or even within (in CBED) the reflections, you can obtain information on the point group symmetry of the structure. To understand this, we first need to go into more detail on these different aspects of symmetry.

3. Point symmetry

Point symmetry operations leave at least one point unchanged. Thus, inherently, they do not have any translation components. There are 10 point symmetry operators, i.e. 1,2,3,4,6, $\bar{1}, m, \bar{3}, \bar{4}, \bar{6}$, the first five X being simple rotations over 360°/X (e.g. 2 is a rotation over 360°/2=180°), m is the mirror plane and the $\bar{X}$ being rotation-inversion axes, first rotating the object over 360°/X and then inverting it. These 10 point symmetry elements can be combined into 32 unique point groups. (Again, see International Tables volume A.)

Point symmetry elements i in a point group can also be represented by "rotation matrices" R_i.

The symbols of the point groups are very systematic, with the sequence of the elements in the symbol corresponding to the main axes per crystal class. An overview of this is given in the Table 1. As an example, in the tetragonal crystal class, the main symmetry elements lie along the c-axis (the fourfold symmetry), a and b (or <a> with <> denoting equivalent directions, and <110>. Thus in any tetragonal point group symbol there will be three positions "**...**" and if you find a 2 on the second position, this means there is a 2-fold axis along the <a> direction.

Crystal class	Characteristic symmetry element	Minimum requirements on the cell parameters	Orientation included in the point group	Point groups belonging to this crystal class
Cubic	4x △	a=b=c α=β=γ=90°	<a> <111> <110>	$\frac{4}{m}\bar{3}\frac{2}{m}$; $\bar{4}3m$; 432 ; $\frac{2}{m}\bar{3}$; 23
Hexagonal	1x ⬡	unique c γ=120° α=β=90° c≠a=b≠c	c <a> <210>	$\frac{6}{m}\frac{2}{m}\frac{2}{m}$; 6mm ; $\bar{6}m2$; 622 ; $\frac{6}{m}$; $\bar{6}$; 6
Trigonal	1x △	unique c γ=120° α=β=90° c≠a=b≠c	c <a>	$\bar{3}\frac{2}{m}$; $3\bar{2}$; 32 ; $\bar{3}$; 3
Tetragonal	1x □	unique c a=b	c <a> <110>	$\frac{4}{m}\frac{2}{m}\frac{2}{m}$; 4mm ; $\bar{4}m2$; 422 ; $\frac{4}{m}$; $\bar{4}$; 4
Orthorhombic	3x 0 of 3x m	a≠b≠c≠a α=β=γ=90°	a b c	$\frac{2}{m}\frac{2}{m}\frac{2}{m}$; mm2 ; 222
Monoclinic	1x 0 of 1x m	unique b α=γ=90° β≠90° a≠b≠c≠a	b	$\frac{2}{m}$; m ; 2
Triclinic	just 1 of $\bar{1}$			$\bar{1}$; 1

Figure 4. Overview of the crystal classes, requirements on the cell parameters, symbol notation for this class and included point groups.

4. Translation symmetry

Besides the translation symmetry that makes each unit cell equivalent to the next, there is also translation symmtry connected to the choice of basis and to certain symmetry operators.
If you choose a centered instead of a primitive unit cell, you will have translation symmetry over vectors within your unit cell. The different centered unit cells are indicated with the letter I, F, A, B, C and R and explained in the Table below.

Table 1. Centered unit cells.

Primitive/ centering	Equivalences between positions	Reflection conditions on hkl
P	x,y,z	none
I	x,y,z; x+1/2, y+1/2, z+1/2	hkl: h+k+l=2n
F	x,y,z; x, y+1/2, z+1/2; x+1/2, y, z+1/2; x+1/2, y+1/2, z	hkl: h+k=2n, k+l=2n, h+l=2n
A	x,y,z; x, y+1/2, z+1/2	hkl: k+l=2n
B	x,y,z; x+1/2, y, z+1/2	hkl: h+l=2n
C	x,y,z; x+1/2, y+1/2, z	hkl: h+k=2n
R	x,y,z; x+1/3, y+1/3, z+1/3	(hexagonal axes): R (obverse) hkl: -h+k+l=3n R (reverse) hkl: h-k+l=3n

A second type of translation symmetry is obtained by combining the point symmetry operators with a translation. This results in screw axes (rotation+translation) and glide planes (mirror+translation). For a screw axis X_n, the translation is always along the direction of axis X, and the length of the translation vector is n/X. For example a 4_1 screw axis combines a 90° rotation with a translation over ¼ of the unit cell. For glide planes, the mirror element can be lying in any orientation a normal mirror plane can be and its orientation is indicated by its place in the symbol (similar to the point group symbols), while the translation vector can be along any combination of unit cell axes parallel to the plane, over half of the unit cell (one fourth is also possible in case of centered cells), the direction of it used as label. So for example a b.. glide plane in an orthorhombic cell has a mirror operation over a plane perpendicular to the a-axis and a translation vector over half of the b-parameter. n would mean a translation over ½ of two axes, d similarly but with a ¼ translation vector.

5. The relation between point symmetry/translation symmetry and reflection conditions

The importance for a TEM user in making a distinction between point symmetry and translation symmetry lies in the fact that the presence of translation symmetry elements causes systematic extinctions, as the translation aspect causes a phase shift between different terms in the structure factor. Point symmetry elements do not cause any extinctions. Thus the point symmetry cannot be derived from analysing which reflections are present/absent.

The lattice centering gives distinct reflection conditions valid for all reflections hkl, and are thus called general conditions (see Table 1). Glide planes give reflection conditions which are only valid for certain zones in reciprocal space and are thus called zonal conditions, and screw axes give reflection conditions only valid along certain directions of reciprocal space and are called serial conditions. The different conditions are tabulated in the International Tables volume A. Thus the translation symmetry can be derived from selected area electron diffraction patterns by determining which reflections are systematically absent.

Point group symmetry can instead be derived using the intensities in for example convergent beam electron diffraction patterns, as will be explained in a dedicated lecture.

6. Bravais lattices

The unit cell centerings that can be combine with the different classes are limited by simple logic. For example, you cannot have an A-centered tetragonal cell, since tetragonality implies that the a and b directions are equivalent, thus that cell should also be B centered. However, for fulfillung lattice requirements then it should also be C centered to make all parallel directions equidistant, i.e. it would become F. However, an F centered tetragonal cell can be reduced to an I-centered tetragonal cell. Thus we eliminate the possibilities A, B and F...and so on. The leftover possible combinations are given in the Table 2 below and are called the Bravais lattices. These are given in the conventional settings. Sometimes researchers publish using non-conventional settings, such as monoclinic I when they have a good reason for this (for example for easy comparison to other structures).

Table 2. Bravais lattices

Crystal Class	Bravais lattices
Triclinic	P
Monoclinic	P, C
Orthorhombic	P, C, I, F
Tetragonal	P, I
Trigonal	P, R
Hexagonal	P
Cubic	P, I, F

7. Space groups

When combining all possible point group elements with translations (centering as well as symmetry elements containing translations), this results in 230 unique space groups. Again, different settings are often possible for one space group. Each space group is derived from the Space Group symmetry of a Bravais lattice, by decreasing the symmetry through replacements of rotation axes and mirror planes by rotation inversion axes or screw axes and glide planes respectively, or just simply eliminating the symmetry element altogether. There are clear group-subgroup relations between the space groups, which go beyond the scope of this short refresher, but are very useful when studying phase transitions. During the lecture, I will explain what is the information contained in the space group tables. This can also be found in the International Tables volume A themselves, or in the teaching edition volume A1. If you do not have the Tables, you have no use for this explanation, therefore I do not include this in this page-limited set of lecture notes.

8. Symmetry of selected area electron diffraction patterns, high resolution transmission electron microscopy images and convergent beam electron diffraction patterns.

Selected area electron diffraction patterns are sections through reciprocal space, not projections, thus the symmetry corresponds to the symmetry of that specific section. It is only two dimensional symmetry and also suffers from Friedel's Law, which adds an inversion centre to any selected area electron diffraction pattern. Therefore reflections hkl are always equivalent in

amplitude with reflections –h-k-l. For example, if you have a section which has 3-fold symmetry in the weighted reciprocal lattice, the addition of the inversion centre will give it 6-fold symmetry in the selected area electron diffraction pattern. Therefore caution needs to be applied before making any conclusions whatsoever straight from selected area electron diffraction patterns about symmetry elements that do not cause systematic extinctions. In the International Tables volume B (a different volume for once), tables can be found with the equivalence relations between the different reflections for each space group. Those tables do not take into account the inversion symmetry caused by the diffraction character, this you need to add yourself. You can also derive these relations using the rotation matrices R_i of the different symmetry elements present in the space group, as h'=hR will have the same amplitude as h.
High resolution transmission electron microscopy images are projections and the symmetry of these images corresponds to the projected symmetry along the zone axis of view. For the main directions, the symmetry of the projections can be found in the International Tables volume A, in the space group tables. These symmetries are plane group symmetries, i.e. contain only symmetry elements staying within a plane. There are only 17 plane groups: p1, p2, pm, pg, cm, p2mm, p2mg, p2gg, c2mm, p4, p4mm, p4gm, p3, p3m1, p31m, p6, p6mm. With the plane groups we talk about mazes instead of unit cells, p is a primitive maze, c a centered maze (in a plane there is only one type of possible centering), 1,2,3,4,6 rotation axes with axis perpendicular to the plane of view, m a mirror plane perpendicular to the plane of view and g a glide plane perpendicular to the plane of view.
Convergent beam electron diffraction patterns contain threedimensional symmetry through the complex dynamic scattering interactions, as will be explained in a different lecture. Point group symmetry can be derived from such patterns.

Quantum Crystallography, Electron Density, and the Kernel Energy Method

Lou Massa
Depts of Chemistry & Physics, Hunter College & the Graduate School, CUNY, New York, USA
lmassa@hunter.cuny.edu

1. Introduction

Here is a suggestion, obvious but also of some importance. Applied to Quantum Crystallography (QCr),[1-13] the kernel energy method (KEM)[10-24] implies the use of the electron density to extract the complete quantum mechanics of a crystallized biomolecule. The possibility of such study is much to be desired, because for truly huge molecular systems it would allow the illumination of important biological problems by using the power of true *ab initio* quantum mechanical explanation.

Most crystal structures are brought to final resolution based upon a model equivalent to a sum of spherical atomic electron densities (the independent atom model (IAM)). That molecular density is:

$$\rho_{\mathrm{IAM}} = \sum_{i=1}^{n} \rho_i^{\mathrm{spherical}} \tag{1}$$

The atoms are positioned so that the crystallographic agreement factor (*R*-factor) is minimized:

$$\delta R = \delta \sum_{\mathbf{K}} w_{\mathbf{K}} [F_{calc.}(\mathbf{K}) - F_{obs.}(\mathbf{K})]^2 = 0 \tag{2}$$

where the structure factor is defined as,

$$F(\mathbf{K}) = \int e^{i\mathbf{K}.\mathbf{r}} \rho(\mathbf{r})\, d^3\mathbf{r}\,, \tag{3}$$

where **K** is the X-ray scattering vector in reciprocal space with components {*h*, *k*, *l*}.

Recognizing that bonding will introduce nonspherical aspects to the molecular density, a better density is obtained by a chemical model which incorporates nonspherical density terms. Thus non-spherical multipole representations of the density have been proposed and in use for over four decades.[25-31] Hansen and Coppens proposed the now oft used model, which given in standard notation reads:[28,29]

$$\rho_{atom} = P_{core}\rho_{core} + P_{valence}\kappa^3 \rho_{valence}(\kappa r) + \sum_{l=0}^{l_{\max}} \kappa'^3 R_l(\kappa' r) \sum_{m=-l}^{l} P_{lm\pm} Y_{lm\pm}(\theta,\varphi)\,, \tag{4}$$

where the total number of electrons of the atom or ion *N* is broken down into sub-populations ($P_{core} + P_{valence} + \sum_{m=0}^{l} P_{lm\pm} = N$), and where ρ_{core} and $\rho_{valence}$ are the normalized densities of the free atom or ion, and R_l are exponential radial functions. The third (valence density) term provides the non-spherical deformation flexibility. This model results in better molecular structures and

better electron densities. High quality crystals and the data collection at low temperatures followed by crystallographic refinement within this model yielded high resolution electron densities that account for the topography of chemical bonding.[32-39]

An example of a multipole deformation density (defined as the difference between a promolecular density obtained by superposing spherical atomic densities and a multipolar density) is given by that shown in Fig. 1. The figure, obtained from Ref. [32], contains contours of a nonspherical deformation density in a peptide plane. Non spherical aspects of the density are evident in this image.

The use of the multipolar representation of the density has become a highly regarded (standard) crystallographic technique.[29,30,33] What more can be asked in the way of improvements? One possible answer would be to extract the complete quantum mechanics from X-ray scattering data.[12,13] Is there a direct connection between X-ray structure scattering data and the quantum mechanics of molecules? Yes, if care is taken to ensure mathematically that the density obtained from crystallography is related to a wavefunction. That relationship is formally called *N*-representability, which, symbolically, can be written as:

$$\rho(1,1') \leftrightarrow \int \Psi^* \Psi d2d3...dN \not\gtrless f(1,1'), \tag{5}$$

where the reader is asked to notice the bijective relation between ρ and Ψ that does not apply for an arbitrary function *f*. Stated differently, a density matrix is *N*-representable if it can be shown to arise from an antisymmetric *N*-body wave function. In particular for:

$$\rho_1 = tr\,\mathbf{P}\boldsymbol{\psi}(\mathbf{r})\boldsymbol{\psi}^{\dagger}(\mathbf{r}), \tag{6}$$

if $\mathbf{P}^2 = \mathbf{P}$, it may be shown that ρ_1 is single determinant *N*-representable.[40] Thus a description for extraction of single determinant quantum mechanics from X-ray scattering may be summarized as follows:

$$\mathbf{P}^2 = \mathbf{P} \tag{7}$$

$$tr\,\mathbf{P} = N \tag{8}$$

$$tr\,\mathbf{Pf}(\mathbf{K}) = F(\mathbf{K}) \tag{9}$$

in which case Eq. (6) delivers a density which is both *N*-representable and consistent with the experimental scattering data.[40]

An example of the theoretical quantum crystallography program just described was an application to the Beryllium crystal.[41] Using simply two basis functions to represent the valence orbital of the Be atom , and the frozen core density from the X-ray tables, the highly accurate data of Larsen and Hansen were used as constraints to fix the single determinant N-representable density.[42] The result was one of high accuracy as measured by the remarkably small *R*-factor of 0.0018. Furthermore, the errors of the density fall exactly within the statistics expected for a random distribution of errors, taking into account the full range of scattering.[41]

[#]Patterned upon a paper submitted to Acta Cryst by Walter Polkosnik, Chérif F. Matta, Lulu Huang, Lou Massa (now under review).

2. A new idea for proceeding to biological molecules

It is seen therefore from the above mentioned beryllium example that true quantum mechanics can indeed be extracted from the X-ray scattering experiment. However, the beryllium "molecule" implies a very small number of electrons. One may ask, what about the case of very large biological molecules? Will exactly the same techniques applied to beryllium serve equally well for thousands or many tens of thousands of atoms, as may occur in biological molecular systems? Can one then, in the same way, find $\mathbf{P}^2 = \mathbf{P}$ that solves the X-ray refinement problem?

The variation of $\mathbf{P}$ against the measured structure factors $F(\mathbf{K})$ becomes impractical as the number of atoms becomes large. For large molecules the number of matrix elements of $\mathbf{P}$ increases with the square of the number of atoms, while the number of X-ray data tend to increase directly with the number of atoms. Clearly there comes a crossover point with increasing numbers of atoms such that insufficient X-ray data are available to deliver unambiguously an *N*-representable density matrix.

A practical way to avoid the dilemma posed by an ever increasing number of atoms is to invoke the Born-Oppenheimer approximation. This asserts that the quantum mechanical electronic structure can be computed at fixed nuclear positions so long as they are known. But the number of X-ray scattering data are sufficient to determine the nuclear positions, so long as the problem is not compounded simultaneously with the many more parameters demanded by a representation of the electron density. Thus it is suggested here to take the nuclear coordinates from the experimental X-ray scattering data and calculating the corresponding electronic structure by the methods of quantum chemistry.[43-46]

Once the problem of extracting quantum mechanics from the X-ray scattering data is divided into experimental determination of the atomic coordinates and the theoretical calculation of the electronic structure, another problem is confronted. And that is that the molecular quantum mechanical computational burden rises as a high power of the number of basis functions used to expand the molecular orbitals. Obviously for large enough molecules, this burden becomes prohibitive. However, another mathematical solution presents itself in the form of the quantum Kernel Energy Method (KEM), which we now discuss.

3. The Kernel Energy Method (KEM)

The KEM makes possible the calculation of true *ab initio* quantum mechanics of large biological molecules of practically any size with high accuracy.[14] This method proceeds by mathematically cutting a large molecule into practicable smaller pieces called "kernels". As an illustrative example, consider the tripeptide Ala-Phe-Thr. This molecule will be broken into three single kernels each a separate amino acid, so the single kernels are Ala, Phe, and Tyr. In order to account for the pairwise interactions between these kernels, calculations are also performed on double kernels. In this case these will consist of two chemically contiguous kernels (Ala-Phe) and (Phe-Thr) and one non-contiguous double kernel (Ala...Thr). The kernels are calculated in the exact geometry as extracted from that of the full target molecule with appropriate capping by

hydrogen atoms to satisfy the dangling bonds (the contributions of these hydrogens cancel in the KEM formulae). Since these kernels are calculated as individual molecules, and importantly, the calculation is inherently parallelizable with possible gigantic saving in computational efficiency.

The kernels are made by cutting across single bonds in the large molecules, and hydrogen caps are added to the kernels to preserve the valence of the atoms at the position of the cuts. As an example for a property that can be predicted from KEM, the total energy of the above tipeptide is reconstructed from the energy of the kernels according to the formula:

$$E_{\underset{(\mathrm{KEM})}{\mathrm{total}}} = \underbrace{\left(E_{\mathrm{Ala}} + E_{\mathrm{Phe}} + E_{\mathrm{Thr}}\right)}_{\text{Energies of kernels}} + \underbrace{\left(\Delta E_{\mathrm{Ala-Phe}} + \Delta E_{\mathrm{Phe-Thr}} + \Delta E_{\mathrm{AlaL\ Thr}}\right)}_{\text{Sum of the energies of pairwise interactions}}, \tag{10}$$

where the first summation is of the energies of the single kernels (the separate amino acids in this case) and the second summation is over all interaction energies of pairs of amino acid residues. The pair-wise interaction energy, say for the first double kernel (Ala-Phe), is written,

$$\Delta E_{\mathrm{Ala-Phe}} = E_{\mathrm{Ala-Phe}} - \left(E_{\mathrm{Ala}} + E_{\mathrm{Phe}}\right). \tag{11}$$

When this last equation is substituted into the expression for the full KEM energy one obtains (in this case):

$$E_{\mathrm{Ala-Phe-Thr}} = \left(E_{\mathrm{Ala-Phe}} + E_{\mathrm{Phe-Thr}} + E_{\mathrm{AlaL\ Thr}}\right) - \left(3-2\right)\left(E_{\mathrm{Ala}} + E_{\mathrm{Phe}} + E_{\mathrm{Thr}}\right), \tag{12}$$

and in general, for a molecule broken into *m* single kernels, the KEM energy is written:

$$E_{\underset{(\mathrm{KEM})}{\mathrm{total}}} = \sum_{a=1}^{m-1}\sum_{b=a+1}^{m} E_{ab} - (m-2)\sum_{c=1}^{m} E_{c} \tag{13}$$

This is the working equation for the KEM.

The KEM approximation has been proposed as a means to obtain fast approximations for electron densities and the underlying and related density matrices.[12,13,24]

$$\rho_2 = \sum_{a=1}^{m-1}\sum_{b=a+1}^{m} \rho_{2(ab)} - (m-2)\sum_{c=1}^{m} \rho_{2(c)}\,, \tag{14}$$

$$\rho_1 = \sum_{a=1}^{m-1}\sum_{b=a+1}^{m} \rho_{1(ab)} - (m-2)\sum_{c=1}^{m} \rho_{1(c)}\,, \tag{15}$$

$$\rho = \sum_{a=1}^{m-1}\sum_{b=a+1}^{m} \rho_{(ab)} - (m-2)\sum_{c=1}^{m} \rho_{(c)}\,. \tag{16}$$

These last three equations represent what may be called the extraction of the complete quantum mechanics from the measured X-ray scattering. The position of the atoms is taken from the crystallography experiment and the electronic structure is calculated by invoking the Born-Oppenheimer approximation followed by the kernel energy method. The quantum mechanics is complete in the sense that all quantum operators of normal interest will have expectation values obtainable from the density and density matrices that have been calculated in the KEM formalism. For example the full molecule electronic energy is of the form

$$E = \left\langle \hat{T}\rho_1 \right\rangle + \left\langle \hat{V}_{ext}\,\rho \right\rangle + \left\langle \hat{V}_{ee}\,\rho_2 \right\rangle \tag{17}$$

which requires ρ_2, ρ_1, and ρ as indicated in the last equation.

Properties of considerable current interest are the total molecular electron density and the total molecular electrostatic potential (ESP) of large macromolecules of biological and nano-technological interest. KEM has been shown to be capable of delivering accurate electron densities sampled at bond and ring critical points (BCPs, and RCPs) in addition to the complete localization-delocalization matrix[47,48] of a graphene nanoribbon.[22] The method has been extended to estimate electric field-induced changes in the properties (*i.e.*, response properties) of a finite graphene flake.[23] There is, thus, strong evidence supporting the confidence in that the extension of KEM to estimate the total molecular electron density and the total molecular electrostatic potential is only awaiting automation, a current interest in our group.

The KEM approximation to the electron density scalar field of a macromolecule can be numerically achieved from cubes of kernel and double kernel electron densities manipulated point-by-point according to Eq. (16). This KEM density can be inserted in the expression for the molecular electrostatic potential[49-54] in which the first term comes from the known molecular structure:

$$V_{\mathrm{KEM}}(\mathbf{r}) = \sum_{\substack{i=1 \\ \mathbf{R}_i \neq \mathbf{r}}}^{\mathrm{M}} \frac{Z_i}{\left|\mathbf{R}_i - \mathbf{r}\right|} - \int \frac{\rho_{\mathrm{KEM}}(\mathbf{r}')}{\left|\mathbf{r}' - \mathbf{r}\right|} d\mathbf{r}' , \tag{18}$$

and where in the first sum the terms with $\mathbf{R}_i = \mathbf{r}$ are eliminated. The reconstructed electrostatic potential ($V_{\mathrm{KEM}}(\mathbf{r})$) can alternatively be obtained from cubes of ESP of single and double kernels manipulated point-wise using an equation of the same form as Eq. (16).

The KEM has been of tested accuracy over a wide range of biological and other large molecules (such as graphene) and has been shown to be accurate and within a range of commonly used chemical models.[14] *Importantly, Eqs. (14-16) imply that the KEM is implementable within any chemical model of choice for the particular problem at hand. It is not inherently restricted to any form of the wavefunction such as a single determinant form.*

For a non-single determinant chemical model, the density matrix ρ_1 of Eq. (15) is expected to be accurate but is not mathematically required to be *N*-representable. This property however can be imposed on the density matrix ρ_1 by casting it into a natural orbital (NO) representation.[55,56] The process for obtaining the natural orbitals involves diagonalizing the density matrix for the full molecule. This is constructed from the "fragment" density matrices of the kernels and double kernels indicated in Eq. (15).[11] The basis set for the full molecule is obtained by collecting together the basis functions from all kernels which constitute the full molecule. If we call the basis set for the full molecule $\boldsymbol{\psi}$, then as in Eq (2) the density matrix for the full molecule is:

$$\rho_1(\mathbf{r}, \mathbf{r}') = 2tr\, \mathbf{R}\, \boldsymbol{\psi}(\mathbf{r}) \boldsymbol{\psi}^{\dagger}(\mathbf{r}') , \tag{19}$$

where $\boldsymbol{\psi}(\mathbf{r})$ is a column vector of atomic orbital basis functions, and where a direct product is implied with its complex conjugate transpose ($\boldsymbol{\psi}\dagger(\mathbf{r}')$). In this last equation $\mathbf{R}$ is the "augmented" matrix[11] constructed by placing the density matrix of the kernels and double kernels into their

appointed positions defined by the full molecule basis $\boldsymbol{\psi}$.

Thus, in an augmented density matrix (ADM), the density matrix of every kernel $\mathbf{r}_i$ or double kernel $\mathbf{r}_{ij}$ (which are all of smaller dimension than for the full molecule density matrix $\mathbf{R}$) is placed within their respective positions as submatrices of the full molecular density matrix:

$$\mathbf{R}_k = \begin{pmatrix} 0 & 0 & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ 0 & 0 & \mathbf{r}_k & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{n\times n}, \tag{20}$$

and

$$\mathbf{R}_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & (\mathbf{r}_{ij})^{11} & 0 & (\mathbf{r}_{ij})^{12} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & (\mathbf{r}_{ij})^{21} & 0 & (\mathbf{r}_{ij})^{22} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{n\times n}, \tag{21}$$

where $\mathbf{r}_k$ is the density matrix of the k^{th} kernel, $\mathbf{R}_k$ is the augmented density matrix of the k^{th} kernel, and where $\mathbf{r}_{ij}$ is the density matrix of the ij^{th} double kernel, $\mathbf{R}_{ij}$ is the augmented density matrix of the ij^{th} double kernel, and where n is the size of the basis set of the full molecule $\boldsymbol{\psi}(\mathbf{r})$. In Eq. (21), an example is shown whereby double kernel matrices denoted by $(\mathbf{r}_{ij})^{AB} = [(\mathbf{r}_{ij})^{BA}]^{\dagger}$ have elements that multiply the bases functions in kernel A by those in kernel B (interaction matrix elements), while $(\mathbf{r}_{ij})^{AA}$ or $(\mathbf{r}_{ij})^{BB}$ multiply basis function within a give kernel (self matrix elements).

The full molecule density matrix $\mathbf{R}$ is approximated from the augmented matrices of the single and double kernels according to the usual KEM formula, that is:

$$\mathbf{R}_{\mathrm{KEM}} \equiv \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \mathbf{R}_{ij}^{aug.} - (m-2)\sum_{k=1}^{m} \mathbf{R}_{k}^{aug.}, \tag{22}$$

where m is the number of single kernels.

The KEM natural orbitals are then obtained from the unitary matrix which diagonalizes $\mathbf{R}_{\mathrm{KEM}}$. The eigenvalues of $\mathbf{R}_{\mathrm{KEM}}$ are the occupation numbers that weight the importance of each natural orbital in the expansion of the density matrix for the full molecule. Thus,

$$\rho_1(\mathbf{r},\mathbf{r}') = \sum_k w_k\, \phi_k^{*}(\mathbf{r}')\,\phi_k(\mathbf{r}) \tag{23}$$

where each ϕ_k is a natural space orbital belonging to an eigenvalue (occupation number) w_k scaled such that they satisfy simultaneously:

$$0 \le w_k \le 2 \tag{24}$$

and,

$$\sum_k w_k = N \tag{25}$$

and where N is the number of occupied natural orbitals.

According to Coleman's theorem[57] the density matrix ρ_1 is N-representable as long as the eigenvalues of the occupied natural orbitals satisfy Eqs. (24) and (25). The corresponding N-representable density ρ is simply the diagonals of ρ_1. Following Müller,[58] a physically useful approximation to ρ_2 can be defined which by integration over the coordinates of electron 2 delivers ρ_1 as is required. The Müller-like construction is:

$$\overline{\rho_2}(1,2,1',2') = \begin{vmatrix} \rho_1(1,1') & \overline{\rho_1}(1,2') \\ \overline{\rho_1}(2,1') & \rho_1(2,2') \end{vmatrix}, \tag{26}$$

where,

$$\overline{\rho_1}(1,1') = \sum_k \sqrt{w_k}\, \phi_k^*(1')\, \phi_k(1). \tag{27}$$

Notice that $\overline{\rho_2}$ is analogous to a single determinant expression but differs in that the off diagonal terms $\overline{\rho_1}$ contain a square root of the eigenvalues, $\sqrt{w_k}$, rather than the eigenvalues themselves, w_k, as occurs in the natural orbital expansion of ρ_1. This however delivers the required form for ρ_1 by integration using the form of $\overline{\rho_2}$ as given by Eq. (26), in which case:

$$\int \overline{\rho_2}(1,2,1',2')\Big|_{2'\to 2}\, d2 = (N-1)\rho_1. \tag{28}$$

4. Conclusions

The principal point of this paper is that a practical way to extract the true quantum mechanics of biological molecules from the X-ray scattering experiment is to invoke the Born-Oppenheimer approximation. In so doing one takes that atomic structure from the X-ray experiment. One recognizes that KEM delivers ρ_2, ρ_1, and ρ and subsequently $F(\mathbf{K})$, the structure factors. Having the X-ray atomic structure allows creating the kernels mathematically and subsequently calculating their *ab initio* electronic structures. Having the electronic structures of the kernels allows the electronic structure of the full molecule to be reconstructed with KEM accuracy according to the its formulas for ρ_2, ρ_1, and ρ [Eqs. (14-16)].

It follows that all expectation values for the full molecule are calculable from these density matrices as shown for the energy, as one example [Eq. (17)]. And finally if the X-ray structure factors $\{F(\mathbf{K})\}$ are calculated with the KEM density $\rho(\mathbf{r})$ [Eq. (3)] that gives the ultimate confirmation of the accuracy of the quantum mechanics obtained from the experiment.

Fig. 2 summarizes the conceptual connection of KEM within the context of quantum crystallography. The figure is to be read starting from the X-rays source that yields the diffraction data that is used to generate the full molecular geometry. The molecule is then partitioned into kernels, sufficiently small to be subjected to quantum chemical computation –

each kernel (or double kernel) separately (and hence highly parallelizable). The calculated density matrices of the kernels are then combined to yield those of the full molecule according to Eqs. (14-16). The KEM density can then be Fourier transformed to deliver calculated structure factors. The calculated and experimental structure factors can be compared by a crystallographic *R*-factor.

Acknowledgments

Funding for this project was provided by the *Natural Sciences and Engineering Research Council of Canada* (NSERC), *Canada Foundation for Innovation* (CFI), and *Mount Saint Vincent University* - (C. F. M.), the *U.S. Naval Research Laboratory* (project # 47203-00 01) and by a *PSC CUNY Award* (project # 63842-00 41) - (L. M.); and the *Office of Naval Research* (ONR) through the *Naval Research Laboratory's Basic Research Program* - (L. H.)

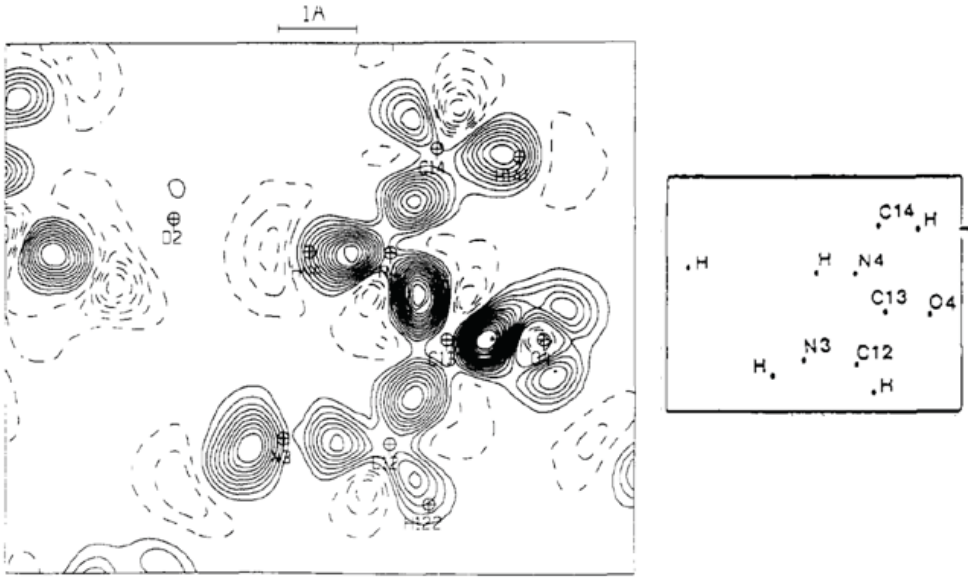


Figure1. Multipolar deformation electron density in the plane of a peptide bond (N4-C13-O4) obtained by Lecomte *et al.* [32] (Left) The experimental density with contour interval of 0.05 $e.Å^{-3}$ and where solid lines indicate positive contours and dashed ones indicate negative contours. The nodal contour has been omitted. (Reproduced with permission of the copyright holder. © 1992 *American Chemical Society*).

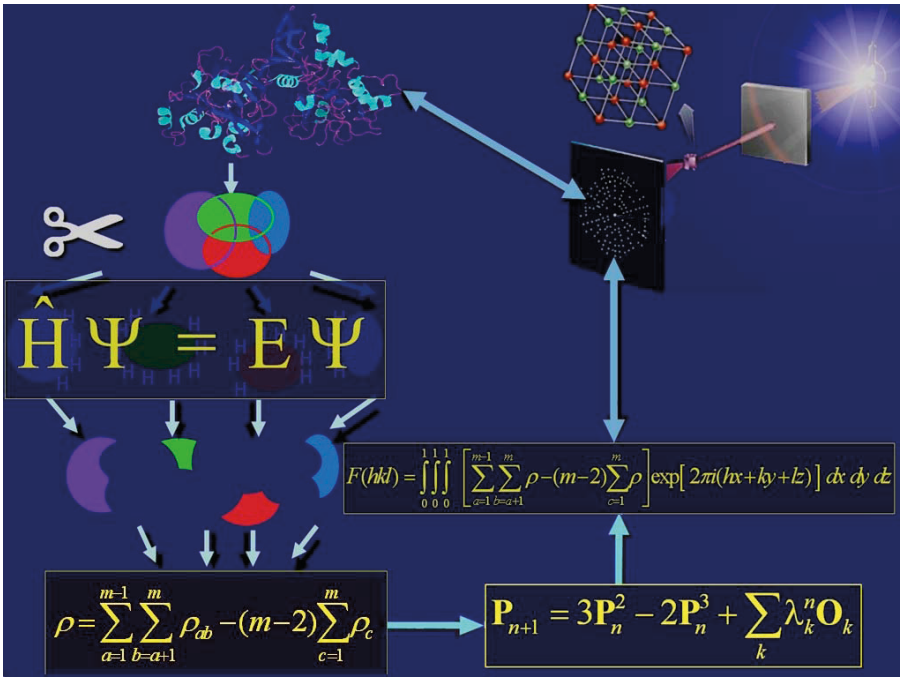


Figure2. A conceptual sketch of the quantum crystallography/kernel energy method (QCr/KEM) scheme from X-rays scattering by the molecule to its fragmentation into pieces passing through the quantum mechanical KEM reconstruction of its density matrices. The reconstructed density and calculated F(K) can then be gauged against experiment. (Reproduced with permission of the copyright holder from Ref. [59] © 2017 *Springer*).

References

(1) Grabowsky, S.; Genoni, A.; Bürgi, H.-B. Quantum crystallography. *Chem. Sci.* 2017, *8*, 4159-4176.

(2) Genoni, A.; Bucinský, L.; Claiser, N.; Contreras-Garcia, J.; Dittrich, B.; Dominiak, P. M.; Espinosa, E.; Gatti, C.; Giannozzi, P.; Gillet, J.-M.; Jayatilaka, D.; Macchi, P.; Madsen, A. Ø.; Massa, L.; Matta, C. F.; Merz Jr., K. M.; Nakashima, P.; Ott, H.; Ryde, U.; Scherer, W.; Schwarz, K.; Sierka, M.; Grabowsky, S. Quantum crystallography: Current developments and future perspectives. *Chem. Eur. J.* 2018, submitted, in review.

(3) Tsirelson, V. Early days of quantum crystallography: A personal account. *J. Comput. Chem.* 2017, in press.

(4) Jayatilaka, D. Wave function for beryllium from X-ray diffraction data. *Phys. Rev. Lett.* 1998, *80*, 798-801.

(5) Jayatilaka, D. *"Using wavefunctions to get more information out of diffraction experiments." In: Modern Charge-Density Analysis, Carlo Gatti & Piero Macchi (Eds.)*; Springer: Berlin, 2012.

(6) Gatti, C.; Macchi, P. *Modern Charge-Density Analysis*; Springer: Berlin, 2012.

(7) Massa, L.; Huang, L.; Karle, J. Quantum crystallography and the use of kernel projector matrices. *Int. J. Quantum. Chem* 1995, *29*, 371-384.

(8) Huang, L.; Massa, L.; Karle, J. Kernel projector matrices for Leu[1]-zervamicin. *Int. J. Quantum. Chem.* 1996, *30*, 479-488.

(9) Huang, L.; Massa, L.; Karle, J. Quantum crystallography, a developing area of computational chemistry extending to macromolecules. *IBM J. Res. & Dev.* 2001, *45*, 409-415.
(10) Huang L.; Massa, L.; Karle, J. "Quantum kernels and quantum crystallography: Applications in biochemistry", in: *Quantum Biochemistry: Electronic Structure and Biological Activity*; Matta C. F. (Ed.), Wiley-VCH: Weinheim, 2010; pp. 3-60.
(11) Polkosnik, W.; Massa, L. Single determinant N-representability and the kernel energy method (KEM) applied to water clusters. *J. Comput. Chem.* 2017, in press.
(12) Massa, L.; Matta, C. F. Exploiting the full quantum crystallography. *Can. J. Chem.* 2018, in press.
(13) Massa, L.; Matta, C. F. Quantum cystallography: A perspective. *J. Comput. Chem.* 2018, in press.
(14) Huang, L.; Massa, L.; Karle, J. Kernel energy method: Application to insulin. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 12690-12693.
(15) Huang, L.; Massa, L.; Karle, J. Kernel energy method: Application to DNA. *Biochemistry* 2005, *44*, 16747-16752.
(16) Huang, L.; Massa, L.; Karle, J. Kernel energy method illustrated with peptides. *Int. J. Quantum Chem.* 2005, *103*, 808-817.
(17) Huang, L.; Massa, L.; Karle, J. The Kernel Energy Method: Application to a tRNA. *Proc. Natl. Acad. Sci. USA* 2006, *103*, 1233-1237.
(18) Huang, L.; Massa, L.; Karle, J. Kernel energy method: Basis functions and quantum methods. *Int. J. Quantum Chem.* 2006, *106*, 447-457 .
(19) Huang, L.; Massa, L.; Karle, J. Kernel energy method applied to vesicular stomatitis virus nucleoprotein. *Proc. Natl. Acad. Sci. USA* 2009, *106*, 1731-1736.
(20) Huang, L.; Bohorquez, H.; Matta, C. F.; Massa, L. The Kernel Energy Method: Application to Graphene and Extended Aromatics. *Int. J. Quantum Chem.* 2011, *111*, 4150-4157.
(21) Huang, L.; Massa, L.; Krupkin, M.; Bashan, A.; Yonath, A. Protoribosome by quantum kernel energy method. *Proc. Natl. Acad. Sci. USA* 2013, *110*, 14900-14905.
(22) Timm, M. J.; Matta, C. F.; Massa, L.; Huang, L. The localization-delocalization matrix and the electron density-weighted connectivity matrix of a finite graphene nanoribbon reconstructed from kernel fragments. *J. Phys. Chem. A* 2014, *118*, 11304-11316.
(23) Huang, L.; Massa, L.; Matta, C. F. A graphene flake under external electric fields reconstructed from field-perturbed kernels. *Carbon* 2014, *76*, 310-320.
(24) Huang, L.; Matta, C. F.; Massa, L. The kernel energy method (KEM) delivers fast and accurate QTAIM electrostatic charge for atoms in large molecules. *Struct. Chem.* 2015, *26*, 1433-1442.
(25) Stewart, R. F. Generalized X-ray scattering factors. *J. Chem. Phys.* 1969, *51*, 4569-4576.
(26) Stewart, R. F.; Bently, J. Generalized X-ray scattering factors in diatomic molecules. *J. Chem. Phys.* 1975, *63*, 3786–3793.
(27) Stewart, R. F. One-electron density functions and many-centered finite multipole expansions. *Isr. J. Chem.* 1977, *16*, 124-131.
(28) Hansen, N. K.; Coppens, P. Testing aspherical atom refinement on small molecules data sets. *Acta Cryst.* 1978, *A34*, 909-921.
(29) Coppens, P. *X-ray Charge Densities and Chemical Bonding*; Oxford University Press, Inc.: New York, 1997.
(30) Tsirelson, V. G.; Ozerov, R. P. *Electron Density and Bonding in Crystals: Principles, Theory and X-ray Diffraction Experiments in Solid State Physics and Chemistry*; Institute of Physics Publishing:

New York, 1996.
(31) Macchi, P.; Gillet, J.-M.; Taulelle, F.; Campo, J.; Claiser, N.; Lecomte, C. Modelling the experimental electron density: only the synergy of various approaches can tackle the new challenges. *Int. U Cryst. J. (IUCrJ)* 2015, *2(Pt 4)*, 441–451.
(32) Pichon-Pesme, V.; Lecomte, C.; Wiest, R.; Bénard, M. Modeling fragments for the ab initio determination of electron density in polypeptides. An experimental and theoretical approach to the electron distribution in Leu-enkephalin trihydrate. *J. Am. Chem. Soc.* 1992, *114*, 2713-2715.
(33) Koritsanszky, T. S.; Coppens, P. Chemical applications of X-ray charge-density analysis. *Chem. Rev.* 2001, *101*, 1583-1628.
(34) Wiest, R.; Pichon-Pesme, V.; Bénard, M.; Lecomte, C. Electron distributions in peptides and related molecules. Experimental and theoretical study of Leu-enkephalin trihydrate. *J. Phys. Chem.* 1994, *98*, 1351-1362.
(35) Espinosa, E.; Lecomte, C.; Molins, E.; Veintemillas, S.; Cousson, A.; Paulus, W. Electron density study of a new non-linear optical material: L-Arginine phosphate monohydrate (LAP). Comparison between *X-X* and *X-(X+N)* refinements. *Acta Cryst.* 1996, *B52*, 519-534.
(36) Jelsch, C.; Pichon-Pesme, V.; Lecomte, C.; Aubry, A. Transferability of multipole charge-density parameters: Application to very high resolution oligopeptide and protein structures. *Acta Cryst.* 1998, *D54*, 1306-1318.
(37) Fernández-Serra, M. V.; Junquera, J.; Jelsch, C.; Lecomte, C.; Artacho, E. Electron density in the peptide bonds of crambin. *Solid State Commun.* 2000, *116*, 395-400.
(38) Lecomte, C.; Guillot, B.; Muzet, N.; Pichon-Pesme, V.; Jelsch, C. Ultra-high-resolution X-ray structure of proteins. *Cell. Mol. Life Sci. (CMLS)* 2004, *61*, 774-782.
(39) Liebschner, D.; Elias, M.; Moniot, S.; Fournier, B.; Scott, K.; Jelsch, C.; Guillot, B.; Lecomte, C.; Chabriere, E. Elucidation of the phosphate binding mode of DING proteins revealed by subangstrom X-ray crystallography. *J. Am. Chem. Soc.* 2009, *131*, 7879-7886.
(40) Clinton, W. L.; Massa, L. J. Determination of the electron density matrix from x-ray diffraction data. *Phys. Rev. Lett.* 1972, *29*, 1363-1366.
(41) Massa, L. ; Goldberg, M.; Frishberg, C.; Boehme, R.; LaPlaca, S. Wave functions derived by quantum modeling of the electron density from coherent X-Ray diffraction: Beryllium metal. *Phys. Rev. Lett.* 1985, *55*, 622-625.
(42) Larsen, F. K.; Hansen, N. K. Diffraction study of the electron density distribution in beryllium metal. *Acta Cryst.* 1984, *B40*, 169-179.
(43) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: New York, 1989.
(44) Levine, I. N. *Quantum Chemistry, (Sixth Edition)*; Pearson Prentice Hall: Upper Saddle River, New Jersey, 2009.
(45) Piela, L. *Ideas of Quantum Chemistry*; Elsevier: Amsterdam, 2007.
(46) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods, (Second Edition)*; Gaussian, Inc.: Pittsburgh, 1996.
(47) Matta, C. F. Molecules as networks: A localization-delocalization matrices approach. *Comput. Theor. Chem.* 2018, *1124*, 1-14.
(48) Matta, C. F. Modeling biophysical and biological properties from the characteristics of the molecular electron density, electron localization and delocalization matrices, and the electrostatic

potential. *J. Comput. Chem.* 2014, *35*, 1165-1198.
(49) Bonaccorsi, R.; Scrocco, E.; Tomasi, J. Molecular SCF calculations for the ground state of some three-membered ring molecules: $(CH_2)_3$, $(CH_2)_2NH$, $(CH_2)_2NH_2^+$, $(CH_2)_2O$, $(CH_2)_2S$, $(CH)_2CH_2$, and N_2CH_2. *J. Chem. Phys.* 1970, *52*, 5270-5284.
(50) Petrongolo, C.; Tomasi, J. The use of electrostatic molecular potential in quantum pharmacology. 1. *Ab initio* results. *Int. J. Quantum Chem.: Quantum Biol. Symp. No. 2* 1975, 181-190.
(51) Bonaccorsi, R.; Scrocco, E.; Tomasi, J. Group contributions to electrostatic molecular potential. *J. Am. Chem. Soc.* 1976, *98*, 4049-4054.
(52) Bonaccorsi, R.; Scrocco, E.; Tomasi, J. Approximate expression of electrostatic molecular potential in terms of completely transferable group contributions. *J. Am. Chem. Soc.* 1977, *99*, 4546-4554.
(53) Tomasi, J. "Use of the electrostatic potential as a guide to understanding molecular properties" in: "*Chemical Applications of Atomic and Molecular Electrostatic Potentials. Reactivity, Structure, Scattering, and Energetics of Organic, Inorganic, and Biological Systems*" Politzer, P.; Truhlar, D. G. (Eds.) New York: Plenum Press; 1981.
(54) Tomasi J.; Cappelli, C.; Mennucci, B.; Cammi, R. "From molecular electrostatic potentials to solvations models and ending with biomolecular photophysical processes" in: *Quantum Biochemistry: Electronic Structure and Biological Activuity (Volume 1)*; Matta, C. F. (Ed.), Wiley-VCH: Weinheim, 2010, pp. 131-170.
(55) Löwdin, P.-O. Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Phys. Rev.* 1955, *97*, 1474-1489.
(56) Löwdin, P.-O.; Shull, H. Natural orbitals in the quantum theory of two-electron systems. *Phys. Rev.* 1956, *101*, 1730-1739.
(57) Coleman, A. J.; Yukalov, V. I. *Reduced Density Matrices: Coulson's Challenge*; Springer: Berlin, 2000.
(58) Müller A. M. K. Explicit approximate relation between reduced two- and one-particle density matrices. *Phys. Lett. A* 1984, *105*, 446-452.
(59) Matta, C. F. A path through quantum crystallography: A short tribute to Professor Lou Massa. *Struct. Chem.* 2017, *28*, 1279-1283.

Accurate X-ray diffraction Measurements

Jacob Overgaard
Department of Chemistry, Aarhus University, Aarhus, Denmark
jacobo@chem.au.dk

1. What characterizes a good data set

In order for us to be able to discuss data quality, as we want to do in this presentation, we first need to agree on what the characteristics of such data are and how we then proceed to measure it.(Herbst-Irmer & Stalke, 2017) We want to have a data set with resolution that extends far beyond the normal values used in your typical structure solution, and this means d-spacing less than 0.5 Å, if at all possible. The aim with doing a charge density determination, which is the primary reason that researchers want to have very accurate data, is to study the chemical bonding.(Koritsanszky & Coppens, 2001, Stalke, 2016) Chemical bonding is strongly associated with the valence electrons, and these are spatially diffuse compared to the compact core electrons, and given the reciprocal connection between the diffraction and physical spaces, one would think that we do not really need high resolution data at all but could do with the low angle data. However, there are at least two main reasons to go to great lengths and measure to very high scattering angle; firstly, we need to separate the thermal smearing of the electron density from the redistribution due to chemical bonding. As will be mentioned later, it is in certain cases possible to obtain complementary information from neutron single crystal diffraction about the atomic displacement parameters, which mathematically describe the thermal smearing, but often this is not an option. Thus, the next best thing is to model the atomic motion using the core electrons alone, which are the sole contributors to the high angle data. Secondly, as you hear in another presentation, the multipole model includes up to 25 additional parameters per atom, so the required number of observations to maintain a ratio of 1:10 for parameters to observations is much higher.

To achieve this we may utilize the strengths of synchrotron radiation and also very low temperatures. We will discuss other needs, as well as the importance of precision and accuracy.

2. Data collection

The collection of accurate data will be the main topic of this talk. Besides, of course, choosing the crystal gives us some freedom, most of us are not able to choose from a large range of instrumental setups but are in fact restricted to the one that for one reason or another has been chosen for your facility, and this is what you have to use. Often your local crystallographer will be a dedicated scientist who has kept the instrument in impeccable working condition, and you will still be able to obtain very good data. However, there is a rapid technological development

these years in X-ray equipment, both in the sources and the detectors, and we will here discuss the different types and their advantages, using wherever possible examples.

a. The crystal

We are faced with many questions when starting a data collection, beginning with the choice of a suitable crystal. There is a saying that in one of several alternative variations goes like this: garbage in, garbage out. For accurate crystallography, this is very true, while standard structure solution is much less sensitive to the data quality. It can however not be stressed enough how important it is to understand that your study of these very tiny details of the charge distribution starts with the choice of crystal. There is no way to get accurate and good data from a poor crystal. However, not all interesting samples are stable in ambient conditions, and it may be necessary to protect the crystals from heat, air or moisture all the way from crystallization glassware to the diffractometer. In the study of unusual chemistry and chemical bonding, compounds are often quite sensitive to the atmosphere and will quickly react and decompose. You can avoid this by mounting crystals in glove box or the X-temp2 cryo-mounting device developed in Göttingen.(Kottke & Stalke, 1993)

b. X-ray sources

Obviously, we need an X-ray source to carry out the experiment. An in-house X-ray source did not change much from its infancy a century ago and until a couple of decades ago; the sealed-tube source. Of course, its lifetime was improved and cost lowered, but the intensity was not really improving. The rotating anode improved the brilliance of the beam (measured in photons/s/area) and they are still very good in-house sources. Not so many years ago the low-powered micro-focus source was developed, which provides high brilliance but running now at about 50 W.

Only recently came then the liquid-metal jet instruments(Hemberg *et al.*, 2003) developed by Excilium in Sweden. As the word suggests they have a liquid stream of metal circulating and the electron beam hits this. A low-melting metal is required as it runs at close to room temperature, so gallium with or without indium are preferred. They provide wavelengths of close to that of Cu for Ga and even shorter for In. However, at the time of writing these notes no charge density experiments have yet been made with these instruments.

The other option is to choose a dedicated synchrotron source. The tunable energy and the massively increased beam-brilliance are the main benefits of applying months in advance and traveling to another country to do an experiment. As the example in this talk suggests, for heavily absorbing samples, using a synchrotron source may in fact be the only way to obtain usable data. However, the use of the synchrotron requires that the other components are optimized for its specific characteristics, such as wavelength variations and extreme countrates. Especially the detector may not be able to cope with this, as we will discuss.

c. X-ray detectors

Most modern diffractometers are equipped with CCD-based detectors, which became widely available two decades ago. It naturally represented a massive improvement for the users to switch from collecting data one Bragg reflection at a time to collecting a large angular region in one go. The CCD detector is fast with readout times of a few seconds. It uses a phosphor material to detect and convert X-ray photons into visible photons. These photons are then detected by the electronics, and stored in individual electron wells until they are read. However, this technology is hampered by a rather high noise level and the overflow of strong intensities into neighbouring pixels. Therefore, the advent of novel pixel-array detectors, which are basically arrays of individual detectors is extremely promising, providing virtually noise free and instantaneous detection.

d. Hydrogen atoms & neutrons

The hydrogen atom constitutes a special problem in the modeling and thus needs special attention. Hydrogen has only one electron, and in most cases it is bonded to a more electronegative atom removing a substantial part of this electron. Therefore, using X-rays alone makes it difficult to determine its exact location and its thermal motion. In many cases, the procedure to overcome this is simply to extend the X-H bond distance to match tabulated values(Allen *et al.*, 1987) obtained from neutron studies, in which the nuclear positions are obtained much more accurately. However, the most accurate is to combine the X-ray data with single crystal neutron diffraction data and perform an X-N refinement,(Coppens, 1967, Blessing, 1995b) in which the pertinent hydrogen parameters are obtained from the neutron data. I will also discuss two more recent alternatives, which are the HAR and the SHADE approach. The former is short for a Hirshfeld Atom Refinement,(Capelli *et al.*, 2014, Jayatilaka & Dittrich, 2008) and involves calculation of aspherical atomic densities based on the Hirshfeld portioning of an ab-initio molecular electron density, Fourier transformation of this to new scattering factors and subsequent refinement. It has been shown to give X-H bond distances highly similar to neutron values, and also in some cases the anisotropic ADPs for hydrogen. On the other hand, SHADE(Madsen & Hoser, 2014, Munshi *et al.*, 2008) uses a TLS-analysis of the X-ray derived non-H ADPs in combination with a library of known internal frequencies for X-H vibrations, and puts this together to give an estimate of hydrogen ADPs.

3. Data reduction

e. Data processing

Once you have collected the data, the next step is to extract the most accurate intensities and their associated errors. Using CCD detectors, the standard has been to perform thin slicing of the reciprocal space and thus have the Bragg reflections covering several adjacent frames. This allows for detailed analysis of the reflection profiles and this is what most software packages by default will do. This works especially well for the weak reflections, which are relatively more affected by the statistical errors. However, it remains to be shown how well this approach works

for the next generation of noise-free detectors, which in combination with synchrotron sources give extremely sharp reflection profiles.

f. Absorption correction

I will also briefly discuss the correction of systematic errors, such as absorption. Most diffractometer software now allows easy indexing of crystal faces, which allows a numerical absorption correction. However, the empirical correction(Blessing, 1995a) may in fact be competitive with the numerical approach.

g. Oblique incidence correction

When diffracted X-rays are detected, it involves the conversion of these high energy photons to either a charge or lower energy photons. The number of charges or photons that are generated in this process depends on the distance the X-rays travel inside the detecting material. It is easy to imagine then that when the incident X-ray is perpendicular to the detector surface, the maximum distance it may travel inside the detector is shorter than if it were entering the detector at an oblique angle. To take this dependency into account an oblique incidence correction needs to be done.(Wu *et al.*, 2002) If uncorrected, the result is that the intensity of the high angle data (if the detector is at 2θ close to zero) will be relatively higher than the low angle data, and thus the ADPs will be much smaller than their true value. The correction depends on the absorption efficiency in the detector at the given energy.

h. Other issues of different importance

We are optimizing all actions that we have control over. In all aspects of this type of experiment, it remains of utmost importance to work with accuracy and precision. The centering of the crystal in the beam, not least when using a small synchrotron beam, is vital. We want to monitor reference frames and hopefully use them to scale the crystal or beam decay. Another sometimes detrimental effect is ice-formation which in principle should be avoidable, but not always is. We may be able to minimize the effect of these effects using post-integration procedures, often incorporated into SADABS(Krause *et al.*, 2015) or equivalent programs.

i. Merging of equivalent reflections and errors

The higher speed of data collection that has been made possible with stronger sources and faster detectors can be used to collect highly redundant data sets. We use these to obtain much better estimates for intensities but also for the associated errors. In the beginning of the 2D detection era, the derived errors from data reduction were often highly underestimated, but these undesirable characteristics have since been cured. Nevertheless, using the reproducibility of intensities by calculating rms values allows a much improved error estimation, and this is often done in programs such as SORTAV. (Blessing, 1987)

4. Post refinement analysis

Having finished the charge density analysis as explained in another talk and workshop, we should take a look at how well the data actually matches the model we have constructed. This

topic has gained some attention in recent years and tools for detailed analysis of errors have been developed.(Zhurov *et al.*, 2008, Henn & Meindl, 2014) This step may in fact reveal some data that are extreme outliers (by not matching the calculated model-based values) and it is necessary to go back to the raw data to find out if there is a plausible explanation.
The host of tools that we use for this are explained at this step in the presentation.

References:

Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1987). *Journal of the Chemical Society, Perkin Transactions 2*, S1-S19.

Blessing, R. H. (1987). *Crystallography Reviews* **1**, 3-58.

Blessing, R. H. (1995a). *Acta crystallographica. Section A, Foundations of crystallography* **51 (Pt 1)**, 33-38.

Blessing, R. H. (1995b). *Acta Crystallogr. Sect. B-Struct. Sci.* **51**, 816-823.

Capelli, S. C., Burgi, H.-B., Dittrich, B., Grabowsky, S. & Jayatilaka, D. (2014). *IUCrJ* **1**, 361-379.

Coppens, P. (1967). *Science* **158**, 1577-1579.

Hemberg, O., Otendal, M. & Hertz, H. M. (2003). *Appl Phys Lett* **83**, 1483-1485.

Henn, J. & Meindl, K. (2014). *Acta Crystallogr. Sect. A* **70**, 248-256.

Herbst-Irmer, R. & Stalke, D. (2017). *Acta Crystallogr. Sect. B* **73**, 531-543.

Jayatilaka, D. & Dittrich, B. (2008). *Acta Crystallogr. Sect. A* **64**, 383-393.

Koritsanszky, T. S. & Coppens, P. (2001). *Chemical Reviews* **101**, 1583-1628.

Kottke, T. & Stalke, D. (1993). *J. Appl. Crystallogr.* **26**, 615-619.

Krause, L., Herbst-Irmer, R., Sheldrick, G. M. & Stalke, D. (2015). *J. Appl. Crystallogr.* **48**, 3-10.

Madsen, A. O. & Hoser, A. A. (2014). *J. Appl. Crystallogr.* **47**, 2100-2104.

Munshi, P., Madsen, A. O., Spackman, M. A., Larsen, S. & Destro, R. (2008). *Acta Crystallogr. Sect. A* **64**, 465-475.

Stalke, D. (2016). *The Chemical Bond I: 100 Years Old and Getting Stronger*, edited by D. M. P. Mingos, pp. 57-88. Cham: Springer International Publishing.

Wu, G., Rodrigues, B. L. & Coppens, P. (2002). *J. Appl. Crystallogr.* **35**, 356-359.

Zhurov, V. V., Zhurova, E. A. & Pinkerton, A. A. (2008). *J. Appl. Crystallogr.* **41**, 340-349.

Dynamics in crystals in the context of quantum crystallography

Anders Ø. Madsen
Department of Pharmacy, University of Copenhagen, Copenhagen, Denmark
a.madsen@sund.ku.dk

Introduction

One of the greatest obstacles when combining quantum mechanical calculations and experimental crystal diffraction data is the modeling of thermal motion. Theoretical estimates of electron densities are based on ab-initio calculations that exclude nuclear motion, but atoms in crystals are always vibrating about their mean positions. Although this motion is lowered as the temperature is reduced, it is never completely absent due to the persistence of zero-point motion. In order to compare theory and experiment it has become practice in most experimental charge density studies to investigate static charge densities, obtained by deconvoluting the nuclear motion. Because most investigators are interested in this static charge density, considerably less focus has been put on the modeling of atomic vibrations. This is a pity for two reasons: first of all, the atomic vibrations are interesting in themselves, and may tell a great deal about the potential energy hypersurface of the atoms and molecules in the crystal and thereby about the thermodynamic and mechanical stability of the crystal; secondly, the models of the static electron density and atomic vibrations are refined against a common set of diffraction data corresponding to the vibrationally averaged density; if either of these models is erroneous, correlations will introduce errors in the other too.
That thermal motion is always present is a fact that must not be disregarded when dealing with an interplay between quantum mechanical models and experimental diffraction data.

Lattice dynamics

Some basic physics of atomic vibrations in crystals is needed in order to explain the models and methods often applied in experimental charge density studies. The theory of lattice dynamics is the basis for deriving the Debye-Waller factor, which describes the reduction of the Bragg diffraction due to atomic motion. The lattice dynamical model can be used to derive Debye-Waller factors from ab-initio calculations, which may then be used in models refined against experimental data – in one way or the other.
A crystal may be considered as one giant molecule. Each atom in the crystal has 3 degrees of freedom. With N unit cells and n atoms in each cell, the crystal has 3nN vibrational degrees of freedom. The vibrations of atoms are correlated and extend throughout the crystal in travelling waves, or phonons. In the theory of lattice dynamics developed by Born and von Kármán, and described in detail in the classical book by Born and Huang [1] the equations of motion of atoms

are set up assuming periodic boundary conditions. Following the theory as outlined in the excellent book by Willis and Pryor [2], the motion of each atom is described by

$$\boldsymbol{u}(kl,t,\boldsymbol{q}) = \boldsymbol{U}(k|\boldsymbol{q})\exp\left[i(\boldsymbol{q}\cdot\boldsymbol{r}(kl) - \omega(\boldsymbol{q})t + \varphi|\boldsymbol{q}))\right] \qquad (1)$$

k labels an atom in the unit cell, and l labels the unit cell. The displacement **u**(kl, t, **q**) of the (k, l) atom from its equilibrium position **r**(kl) depends on the wave vector **q** of the travelling wave. The displacement vector **U**(kj**q**) describes the maximum amplitude and the direction of motion of the (k, l) atom, as produced by the travelling wave of wave vector **q**. **U** is independent of the unit cell l, because the motion of equivalent atoms (k) in different cells (l) have identical amplitude and direction and differ only in phase. This is *Bloch's theorem*, and introduces an enormous simplification as it allows us to restrict attention to the 3n equations of motion of the n atoms of just one cell, rather than the equations of motion of the 3nN atoms in the crystal. The frequency ω in Eq. 1 is a continuous function of **q**, and the dependence of ω on **q** is called the *dispersion relation* for the propagation direction defined by **q**. For a given **q**, Eq. 1 describes 3n modes of vibration, so that we need N different q-vectors to describe the atomic vibrations of the crystal. These **q** vectors can be chosen to be uniformly distributed within the so-called first Brillouin zone, or simply *the* Brillouin zone. This is the region, centered on the origin of reciprocal space and bound by planes drawn as perpendicular bisectors of the vectors joining the origin to the nearest reciprocal lattice points.

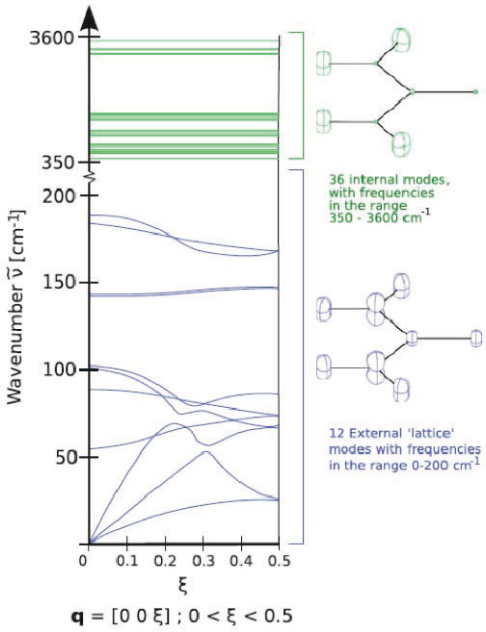


Figure 1. Dispersion relations of Urea - and corresponding mean square displacements. Figure from the book 'Modern charge density analysis', ed. Macchi and Gatti.

Figure 1 shows dispersion relations for the 3 x 16 = 48 modes of vibration for urea, which has n=16 atoms in the unit cell. The dispersion relations are shown in the first Brillouin zone in the direction of the **q** vector [0 0 ξ], $0 < \xi < 0.5$. The dispersion curves are based on a force-constant model fitted against inelastic neutron scattering measurements [3] and ab-initio calculations [4]. Once the normal modes and frequencies have been obtained from a force-field model, they can be used to calculate the vibrational contributions to the thermodynamics of the crystal, i.e. the entropy and heat capacity. They can also be used to calculate the atomic mean square displacement tensors, and thereby the Anisotropic Displacement Parameters (ADPs) which are normally part of crystallographic models. The mean square displacement tensor of a vibrating atom k, may be written in terms of a summation of contributions from all the 3nN normal modes of vibration:

$$\boldsymbol{B}_{atom}(k) = \frac{1}{N m_k} \sum_{j\boldsymbol{q}} \frac{E_j(\boldsymbol{q})}{\omega_j^2(\boldsymbol{q})} e(k|j\boldsymbol{q}) \boldsymbol{e}^*((k|j\boldsymbol{q}))^T \qquad (2)$$

where $\mathbf{e}(k|j\mathbf{q})$ represents the kth component of a normalized complex eigenvector $\mathbf{e}(j\mathbf{q})$, and corresponds to atom k in normal mode j along the wavevector **q**. ω_j is the frequency of mode j, m_k is the mass of atom k, and $E_j(\mathbf{q})$ is the energy of the mode, given by

$$E_j(\boldsymbol{q}) = \hbar\omega_j(\boldsymbol{q}) \left(\frac{1}{2} + \frac{1}{\exp(\hbar\omega_j(\boldsymbol{q})/k_B T) - 1} \right) \qquad (3)$$

$\mathbf{B}_{atom}(k)$ in Eq. 2 is the atomic mean square displacement tensor. It is a symmetric 3 x 3 tensor equivalent to the tensor of ADPs, as further described below. The normal mode vectors $\mathbf{e}(j\mathbf{q})$ derived from the lattice dynamical model contain information about the correlation of atomic motion. This correlation is not directly available from a single temperature experiment, but some information can be derived using a rigid-body approximation and via multi-temperature experiments, *vide infra*.

Atomic displacement parameters

We have sketched how the theory of lattice dynamics can be used to describe the atomic vibrations in a crystal in terms of normal modes of vibration. The usual models adopted for structure refinement and charge-density analysis in crystallography do not use a lattice dynamical approach, but consider the atoms as individual harmonic oscillators, moving in the mean field of the surrounding atoms. In this approach, the mean thermal electron density of an atom is considered to be the convolution of a static density $\rho_k(\mathbf{r})$ with the probability density function $p_k(\mathbf{r})$ describing the probability of having atom k displaced from its reference position $\mathbf{r}_{k0}$;

$$\langle \rho_k(\boldsymbol{r}) \rangle = \int \rho_k(\boldsymbol{r} - \boldsymbol{r}_k)\, p_k(\boldsymbol{r}_k - \boldsymbol{r}_{k0}) d\boldsymbol{r}_k \,. \qquad (4)$$

It is important to realize that in this model, the static atomic electron density $\rho_k(\mathbf{r})$ is not deformable: As the atom is vibrating and thereby displaced from its equilibrium position, the electron density is rigidly following the nucleus. This approximation is called the rigid pseudo

atom approximation. Whether this approximation is reasonable is difficult to investigate. The question is whether the perturbation of the atomic electron density caused by (vibrational) changes in the internuclear distances is of a magnitude that can be detected by experimental charge density studies. At least, the model is sufficient to provide displacement parameters in excellent agreement with parameters derived from inelastic neutron scattering measurements [5,6].
The X-ray structure factor for the scattering vector **h** is given by the Fourier transform of the average electron density of the unit cell,

$$F(\boldsymbol{h}) = \int \langle \rho(\boldsymbol{r}) \rangle e^{2\pi i \boldsymbol{h}\cdot\boldsymbol{r}}\, d\boldsymbol{r}. \qquad (5)$$

Each atomic contribution consists of a form factor $f_k(\mathbf{h})$, which is the Fourier transform of the static density $\rho_k(r)$, multiplied by a term that is the Fourier transform of the atomic probability density function

$$F(\boldsymbol{h}) \approx \sum_{k=1}^{N} n_k\, f_k(\boldsymbol{h})\, T_k(\boldsymbol{h}) \exp\,(2\pi i \boldsymbol{h}\cdot\boldsymbol{r}_{k0}). \qquad (6)$$

The scattering factor, or atomic form factor, of atom *k* is

$$f_k(\boldsymbol{h}) = \int \rho_k(\boldsymbol{r}) \exp(2\pi i\, \boldsymbol{h}\cdot\boldsymbol{v})\, d\boldsymbol{v} \qquad (7)$$

Where v is a positional vector with origin at atom k. Likewise the Debye Waller factor is the Fourier transform of the atomic probability density function $p_k(\mathbf{u})$;

$$T_k(\boldsymbol{h}) = \int p_k(\boldsymbol{u}) \exp(2\pi i\, \boldsymbol{h}\cdot\boldsymbol{u})\, d\boldsymbol{u} \qquad (8)$$

This last term contains the dependence of the structure factor on the atomic displacement. In most cases the pdf is approximated as an anisotropic Gaussian function. Under such circumstances, T can be written as

$$T = \exp\left(-2\pi^2 \sum_{j=1}^{3}\sum_{l=1}^{3} h_j a^j U^{jl} a^l h_l\right) \qquad (9)$$

The component U^{jl} is one form of the anisotropic displacement parameters; they can be described in other bases as well. The parameters have dimension $(\text{length})^2$ and can be directly associated with the mean square displacements of the atom considered in the corresponding directions. For a thorough description of displacement parameters read the paper by Trueblood et al [7].

Validation and analysis of atomic motion

How do we know whether a set of ADPs are sound? Do they reflect the vibrations in the crystal, or are there artifacts due to e.g. static disorder in the crystal, or due to experimental or modeling errors? For a discussion of commonly found problems, seen in the light of differences between structures from X-ray and neutron experiments, read Blessing's paper [8].
For the purpose of structural refinement an inspection of the equal-probability ellipsoids (Fig. 2) often suffices to elucidate problems. Abnormally oblate or prolate ellipsoids are usually caused by static disorder, where several conformations of molecules or parts of molecules occur in otherwise identical unit cells. Static displacive disorder is quite common in crystals, however very few systems showing obvious signs of static disorder have been subjected to a quantum-

crystallographic studies. Static disorder is temperature independent, as opposed to the thermal vibrations, and multi-temperature studies are an obvious way of distinguishing between the different types of contributions to the ADPs. Ellipsoids that are elongated in the same direction for all atoms are typically caused by a missing or erroneous absorption correction. It is a necessary but not sufficient condition that the ADPs fulfill the rigid bond test proposed by Hirshfeld [9]; covalently bonded atoms of similar mass (e.g. second-row atoms in organic molecules) must have similar mean-square displacements in the direction of the bond. The mean square displacement of an atom k in the direction of the unit vector v is given by $\mathbf{v}\,\mathbf{B}(k)\,\mathbf{v}^T$ where $\mathbf{B}(k)$ is the mean square displacement tensor (in Cartesian representation) of atom k, and $\mathbf{v}^T$ is the transpose of $\mathbf{v}$. Differences of more than 10^{-4} Å^2 should be viewed with skepticism. The rigid bond test can be extended to include non-bonded atoms in the structure [10]. If the mean square displacements between non-bonded atoms in a molecule are found to obey the rigid bond test, the entire molecule is probably vibrating as a rigid body. It may also be found that only a part of the molecule is moving as a rigid body. In either case, the ADPs of the structure can be subjected to a rigid body analysis or segmented rigid body analysis, as explained further below.

Rigid body analysis

Rigid body analysis is an attempt to analyze the atomic mean square displacements of a molecule as if the molecule was vibrating as a rigid unit, independent of the motion of the surrounding molecules in the crystal.

Following the pioneering work of Cruickshank [11,12] researchers have analyzed the ADPs as if they originated from collective motion with a considerable amount of success [19–24]. The most well-known model is the Translation/Libration/Screw (TLS) model developed by Schomaker and Trueblood [13]. The ADPs do not contain information about the correlation of motion between different atoms – however since the energy of the molecular modes depend on the temperature via Eq. 2 and 3, multitemperature experiments can recover part of this correlation, as shown by Bürgi and co-workers [14,15].

A range of computer programs have been developed to perform rigid body analysis, either using the TLS formalism (PLATON [16] and THMA11 [13]) or related models (EKRT [17,18] and NKA [15]). A good review of the TLS method is given by Dunitz et al [19].

Estimating ADPs for Hydrogen Atoms

Hydrogen atoms are difficult to handle, seen from the point of view of the X-ray crystallographer. The scattering from hydrogen is very weak because of the low electron density, and the density is also polarized towards the covalent bond. If possible, it is therefore advantageous to perform a complimentary neutron diffraction study. However, the limited access to neutron facilities, and the problems associated with growing very large crystals (several mm^3) makes this impractical in many cases. A number of ways to estimate the ADPs have therefore been proposed.

The SHADE approach

It is possible to analyze the vibrational motion of hydrogen atoms in a similar vein as the statistical analysis of X–H bond lengths derived from neutron diffraction studies found in International Tables for Crystallography [20]. When the total atomic mean square displacement tensor **U** has been determined from neutron diffraction experiments, and the rigid molecular motion $\mathbf{U}_{\text{rigid}}$ has been determined from a rigid-body analysis of the non-hydrogen ADPs, it becomes possible to get an estimate of the internal motion of the hydrogen atoms;

$$U^{ij} = U^{ij}_{internal} + U^{ij}_{rigid}$$

It was noted by Johnson [21] that the mean square displacements derived from $\mathbf{U}_{\text{internal}}$ of hydrogen atoms was in good agreement with spectroscopic information, showing systematic trends corresponding to the functional group that hydrogen was part of. Similar observations were done by Craven and co-workers in the analysis of several systems[22–25]. The internal torsional motion of a range of librating groups, including methyl, carboxyl and amino groups was also thoroughly investigated by Trueblood and Dunitz [26] based on more than 125 neutron diffraction studies of molecular crystals from the literature.

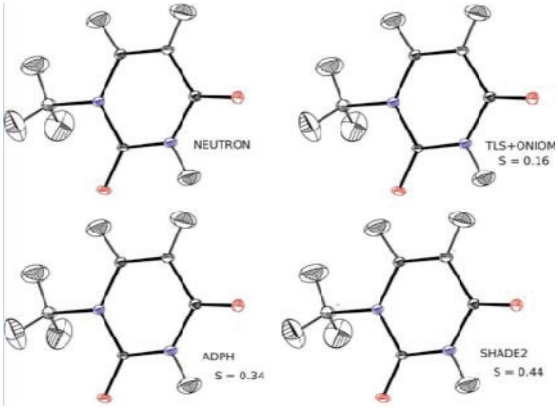


Figure 2. Estimated H atom ADPS for 1-methyl uracil using various approaches. Reproduced from Munshi et al.

Inspired by these results we analyzed a range of neutron structures found in the literature, and the estimates of internal motion were collected in a 'library' and later improved and enhanced with more statistical material [27,28]. The present SHADE2 library provides mean values of internal stretch modes as well as in-plane and out-of-plane bending modes for a range of chemical groups involving hydrogen bound to C, N and O. The library forms the basis for assigning anisotropic displacement parameters to hydrogen atoms in the SHADE server [29] which allows users to submit a CIF file containing the atomic coordinates and the ADPs of the non-hydrogen atoms. The server performs a TLS analysis using the THMA11 program, and combines the rigid body motion with the internal motion obtained from analysis of neutron diffraction data. It is possible to perform a segmented rigid body analysis using the attached rigid group approach of the THMA11 program. The SHADE server is available at the web-address http://shade.ki.ku.dk.

Whereas SHADE uses information from neutron diffraction studies, ADPH uses spectroscopic information [30], and TLS+ONIOM derives the internal motion from ab-initio ONIOM calculations [31]. Lübben et al [32] have developed the TLS+ONIOM idea further and provides new TLS software and a library of ab-initio computed internal vibrations. A related approach is the SHADE3 approach, where the internal modes are derived from periodic DFT [33].
The ADPH, SHADE and TLS+ONIOM approaches have been compared by Munshi et al. [28]. They differ primarily in the way the internal motion is estimated. The ADPs of hydrogen atoms in 1-methyl-uracil based on these approaches are compared in Fig. 2. The mean similarity index (see the paper by Munshi et al for further details) is given on the figure. All models are in excellent agreement with the ADPs based on neutron diffraction experiments, and this was also the general conclusion in the comparison by Munshi et al., where the SHADE server was recommended as a routine procedure for deriving estimates of H-atom ADPs suitable for charge-density studies of molecular crystals.

Atomic motion derived from Force-field or ab-initio calculations

In the late 1970s and 1980s Gramaccioli and co-workers made important progress in the evaluation of temperature factors based on Born–von Kármán lattice-dynamical force field calculations [34]. These studies showed a reasonable agreement with experimental temperature factors from neutron diffraction studies. In recent years, it has become feasible to derive the force constants for lattice-dynamical calculations of vibrational modes and thermodynamic properties from periodic DFT (density functional theory) calculations, and these types of calculations have been used intensively to investigate the vibrational and thermodynamic properties of mostly metals and inorganic materials [35]. There are well established methods for calculating mean square displacements (MSDs) based on force-field calculations. Only recently have these methods been extended to periodic DFT calculations on molecular crystals, because the calculation of ADPs for molecular crystals has additional challenges that are not found for extended solids. Whereas the dominant forces between the atoms in extended solids are strong and of covalent or ionic character, molecular crystals are held together by much weaker forces. In fact, dispersion forces may be dominant in systems without hydrogen-bonding capabilities. In these situations, current ab-initio methods for solids, in particular the widely used DFT methods, may prove to be unable to quantitatively describe the dynamics of the system.

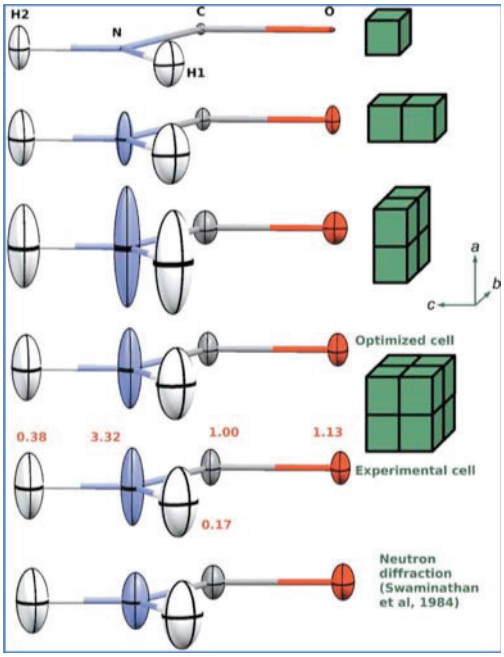


Figure 3. Ellipsoids of thermal motion for the Urea crystal estimated from periodic DFT calculations. Figure from Madsen et al (2013).

Morrison and co-workers [36,37] have used Car–Parrinello molecular dynamics (MD) simulations to investigate the anharmonic atomic motion in crystals. Based on their calculations, they suggest new models of anharmonic motion that will reduce the number of parameters typically used for these models (the Gram–Charlier expansion). Nemkevich et al. [38] used traditional force-field-based MD simulations to obtain isotropic harmonic motion. Comparing their results with experimental results, they generally observe that the computed MSDs are much smaller than the experimentally observed ones.

Dittrich et al. [39] have used ONIOM (quantum mechanical/molecular mechanics hybrid) calculations to obtain ADPs for a number of molecular crystals. Their results compare well with experiments at ultra-low temperatures (10–20 K) when the computed ADPs are scaled against the observed ADPs.

Madsen et al [4] used DFT methods in combination with an empirical dispersion term, proposed by Grimme [40] and tested for crystalline systems [41]. This approach was tested on the urea crystal, see Fig. 3. The results were promising, although not perfect. Since the work of Madsen et al (2013) the group of Dronskowski has published a range of papers on obtaining ADPs for molecular crystals from dispersion-corrected periodic DFT (e.g. Deringer *et al.*, 2014; George *et al.*, 2015). They have a website where more information and help can be found (www.ellipsoids.de).

A new approach: Dynamic Quantum Crystallography

Information on the correlation of atomic motion is lost in the standard elastic scattering experiment. However, as we have discussed above, the temporal and spatial average of the atomic fluctuations – mean square displacements - can be retrieved from a diffraction measurement. In recent work we have combined this information with lattice-dynamical models derived from periodic DFT calculations [44,45]. In this approach, the amplitudes of the acoustic and lowest-frequency optical phonons are refined against the diffraction intensities. In our simplest model, these phonon modes are approximated by the motion at the Gamma point of the Brillouin zone.

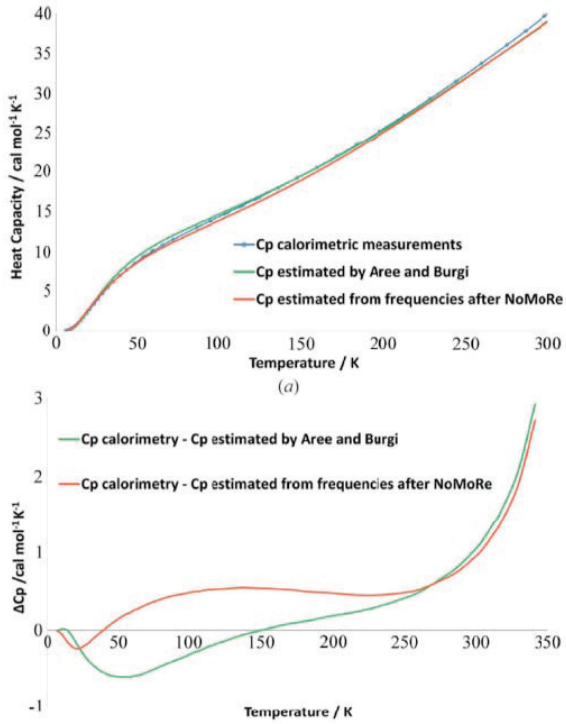


Figure 4. Top: Heat capacity of Naphthalene (Cp): from thermodynamic measurements (blue curve), calculated from frequencies obtained after NoMoRe (red curve) and estimated by Aree and Bürgi (green curve). Bottom:) difference between Cp from calorimetric measurements and from the models. Figure from Hoser & Madsen (2017).

Despite the very simple lattice dynamical model, these Normal Mode Refinements (NoMoRe) captures essential information about the crystal dynamics from the experiments. In figure 4 we compare the heat capacity of naphthalene obtained from calorimetric measurements against the heat capacities obtained by the NoMoRe procedure, as well as with the related models of Bürgi and Aree [46].

The atomic mean square displacements obtained by fitting the normal modes against the diffraction intensities compare well with the displacements obtained from standard crystallographic models, and additionally the hydrogen atom anisotropic displacements compare

well with independent information from neutron diffraction experiments. However, by combining aspherical atom refinement and normal mode refinement it is evident that there is information in the diffraction experiments that is not captured by the model (Sovago, Hoser, Madsen in preparation); there is plenty of room for improvements.

Diffuse scattering

One obvious next step in quantum crystallographic studies of dynamics is to model *thermal diffuse scattering (TDS).* Studies of diffuse scattering from crystals are experiencing a renaissance in these years. This is due to the advent of very sensitive low-noise detectors and because high-performance computing has made it possible to construct ab-initio models of the crystals which can explain the diffuse patterns. Diffuse scattering patterns originate from ordering at length scales larger than the unit cell dimensions. The ordering can be of either static or dynamic character; and in both cases reveal important information about the physical properties of the crystal.

The thermal diffuse scattering (TDS) signal can be diminished by cooling the crystals to very low temperatures, but it can never be fully removed. If TDS is not accounted for it will give rise to additional systematic changes in the Bragg intensities, and thus create artifacts in the crystallographic models, as we have recently demonstrated in a model study on silicon and cubic boron nitride [47]. This implies that even in quantum crystallographic studies where the dynamics is of secondary interest, it is important to have an accurate model of motion in order to properly take these contributions into account.

(1) Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices*; Oxford Classic Texts in the Physical Sciences; Oxford University Press: Oxford, New York, 1998.

(2) Willis, B. T. M.; Pryor, A. W. *Thermal Vibrations in Crystallography*; Cambridge University: Cambridge, 1975.

(3) Lefebvre, J.; More, M.; Fouret, P.; Hennion, B.; Currat, R. Lattice Vibrations in Deuterated Urea. *J. Phys. C Solid State Phys.* **1975**, *8* (13), 2011.

(4) Madsen, A. Ø.; Civalleri, B.; Ferrabone, M.; Pascale, F.; Erba, A. Anisotropic Displacement Parameters for Molecular Crystals from Periodic Hartree–Fock and Density Functional Theory Calculations. *Acta Crystallogr. A* **2013**, *69* (3), 309–321.

(5) Flensburg, C.; Stewart, R. F. Lattice Dynamical Debye-Waller Factor for Silicon. *Phys. Rev. B* **1999**, *60* (1), 284.

(6) Willis, B. T. M.; Howard, J. a. K. Do the Ellipsoids of Thermal Vibration Mean Anything? – Analysis of Neutron Diffraction Measurements on Hexamethylenetetramine. *Acta Crystallogr. A* **1975**, *31* (4), 514–520.

(7) Trueblood, K. N.; Bürgi, H. B.; Burzlaff, H.; Dunitz, J. D.; Gramaccioli, C. M.; Schulz, H. H.; Shmueli, U.; Abrahams, S. C. Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. *Acta Crystallogr. A* **1996**, *52* (5), 770–781.

(8) Blessing, R. H. On the Differences between X-Ray and Neutron Thermal Vibration Parameters. *Acta Crystallogr. B* **1995**, *51* (5), 816–823.
(9) Hirshfeld, F. L. Can X-Ray Data Distinguish Bonding Effects from Vibrational Smearing? *Acta Crystallogr. A* **1976**, *32* (2), 239–244.
(10) Rosenfield, R. E.; Trueblood, K. N.; Dunitz, J. D. A Test for Rigid-Body Vibrations Based on a Generalization of Hirshfeld's `rigid-Bond' Postulate. *Acta Crystallogr. Sect. A* **1978**, *34* (5), 828–829.
(11) Cruickshank, D. W. J. The Determination of the Anisotropic Thermal Motion of Atoms in Crystals. *Acta Crystallogr.* **1956**, *9* (9), 747–753.
(12) Cruickshank, D. W. J. The Analysis of the Anisotropic Thermal Motion of Molecules in Crystals. *Acta Crystallogr.* **1956**, *9* (9), 754–756.
(13) Schomaker, V.; Trueblood, K. N. On the Rigid-Body Motion of Molecules in Crystals. *Acta Crystallogr. B* **1968**, *24* (1), 63–76.
(14) Bürgi, H. B.; Capelli, S. C. Dynamics of Molecules in Crystals from Multi-Temperature Anisotropic Displacement Parameters. I. Theory. *Acta Crystallogr. A* **2000**, *56* (5), 403–412.
(15) Capelli, S. C.; Förtsch, M.; Bürgi, H. B. Dynamics of Molecules in Crystals from Multi-Temperature Anisotropic Displacement Parameters. II. Application to Benzene (C_6D_6) and Urea [$OC(NH)_2$]. *Acta Crystallogr. A* **2000**, *56* (5), 413–424.
(16) Spek, A. L. PLATON, An Integrated Tool for the Analysis of the Results of a Single Crystal Structure Determination. *Acta Crystallogr. A* **1990**, *46* (s1), 34–34.
(17) He, X. M.; Craven, B. M. Internal Molecular Vibrations from Crystal Diffraction Data by Quasinormal Mode Analysis. *Acta Crystallogr. A* **1985**, *41* (3), 244–251.
(18) He, X.-M.; Craven, B. M. Internal Vibrations of a Molecule Consisting of Rigid Segments. I. Non-Interacting Internal Vibrations. *Acta Crystallogr. A* **1993**, *49* (1), 10–22.
(19) Dunitz, J. D.; Maverick, E. F.; Trueblood, K. N. Atomic Motions in Molecular Crystals from Diffraction Measurements. *Angew. Chem. Int. Ed. Engl.* **1988**, *27* (7), 880–895.
(20) Allen, F. H.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. Typical Interatomic Distances: Organic Compounds. In *International Tables for Crystallography Volume C: Mathematical, physical and chemical tables*; Prince, E., Ed.; International Tables for Crystallography; Springer Netherlands, 2006; pp 790–811.
(21) Johnson, C. K. Generalized Treatments for Thermal Motion. In *Thermal Neutron Diffraction*; Willis, B. T. M., Ed.; Oxford University Press, 1970; pp 132–160.
(22) Gao, Q.; Weber, H. P.; Craven, B. M.; McMullan, R. K. Structure of Suberic Acid at 18.4, 75 and 123 K from Neutron Diffraction Data. *Acta Crystallogr. B* **1994**, *50 (Pt 6)*, 695–703.
(23) Kampermann, S. P.; Sabine, T. M.; Craven, B. M.; McMullan, R. K. Hexamethylenetetramine: Extinction and Thermal Vibrations from Neutron Diffraction at Six Temperatures. *Acta Crystallogr. A* **1995**, *51* (4), 489–497.
(24) Luo, J.; Ruble, J. R.; Craven, B. M.; McMullan, R. K. Effects of H/D Substitution on Thermal Vibrations in Piperazinium Hexanoate-H11,D11. *Acta Crystallogr. B* **1996**, *52* (2), 357–368.
(25) Weber, H.-P.; Craven, B. M.; Sawzik, P.; McMullan, R. K. Crystal Structure and Thermal Vibrations of Cholesteryl Acetate from Neutron Diffraction at 123 and 20 K. *Acta Crystallogr. B* **1991**, *47* (1), 116–127.

(26) Trueblood, K. N.; Dunitz, J. D. Internal Molecular Motions in Crystals. The Estimation of Force Constants, Frequencies and Barriers from Diffraction Data. A Feasibility Study. *Acta Crystallogr. B* **1983**, *39* (1), 120–133.

(27) Madsen, A. Ø.; Mason, S.; Larsen, S. A Neutron Diffraction Study of Xylitol: Derivation of Mean Square Internal Vibrations for H Atoms from a Rigid-Body Description. *Acta Crystallogr. B* **2003**, *59* (5), 653–663.

(28) Munshi, P.; Madsen, A. Ø.; Spackman, M. A.; Larsen, S.; Destro, R. Estimated H-Atom Anisotropic Displacement Parameters: A Comparison between Different Methods and with Neutron Diffraction Results. *Acta Crystallogr. A* **2008**, *64* (4), 465–475.

(29) Madsen, A. Ø. *SHADE* Web Server for Estimation of Hydrogen Anisotropic Displacement Parameters. *J. Appl. Crystallogr.* **2006**, *39* (5), 757–758.

(30) Roversi, P.; Destro, R. Approximate Anisotropic Displacement Parameters for H Atoms in Molecular Crystals. *Chem. Phys. Lett.* **2004**, *386* (4–6), 472–478.

(31) Whitten, A. E.; Spackman, M. A. Anisotropic Displacement Parameters for H Atoms Using an ONIOM Approach. *Acta Crystallogr. Sect. B* **2006**, *62* (5), 875–888.

(32) Lübben, J.; Bourhis, L. J.; Dittrich, B. Estimating Temperature-Dependent Anisotropic Hydrogen Displacements with the Invariom Database and a New Segmented Rigid-Body Analysis Program. *J. Appl. Crystallogr.* **2015**, *48* (6), 1785–1793.

(33) Madsen, A. Ø.; Hoser, A. A. SHADE3 Server: A Streamlined Approach to Estimate H-Atom Anisotropic Displacement Parameters Using Periodic Ab Initio Calculations or Experimental Information. *J. Appl. Crystallogr.* **2014**, *47* (6), 2100–2104.

(34) Gramaccioli, C. M.; Filippini, G. Lattice-Dynamical Evaluation of Temperature Factors in Non-Rigid Molecular Crystals: A First Application to Aromatic Hydrocarbons. *Acta Crystallogr. A* **1983**, *39* (5), 784–791.

(35) Baroni, S. Phonons and Related Crystal Properties from Density-Functional Perturbation Theory. *Rev. Mod. Phys.* **2001**, *73* (2), 515–562.

(36) Reilly, A. M.; Wann, D. A.; Gutmann, M. J.; Jura, M.; Morrison, C. A.; Rankin, D. W. H. Predicting Anisotropic Displacement Parameters Using Molecular Dynamics: Density Functional Theory plus Dispersion Modelling of Thermal Motion in Benzophenone. *J. Appl. Crystallogr.* **2013**, *46* (3), 656–662.

(37) Reilly, A. M.; Morrison, C. A.; Rankin, D. W. H. Using Molecular-Dynamics Simulations to Understand and Improve the Treatment of Anharmonic Vibrations. I. Study of Positional Parameters. *Acta Crystallogr. A* **2011**, *67* (4), 336–345.

(38) Nemkevich, A.; Bürgi, H.-B.; Spackman, M. A.; Corry, B. Molecular Dynamics Simulations of Structure and Dynamics of Organic Molecular Crystals. *Phys. Chem. Chem. Phys.* **2010**, *12* (45), 14916.

(39) Dittrich, B.; Pfitzenreuter, S.; Hübschle, C. B. On QM/MM and MO/MO Cluster Calculations of All-Atom Anisotropic Displacement Parameters for Molecules in Crystal Structures. *Acta Crystallogr. A* **2012**, *68* (1), 110–116.

(40) Grimme, S. Accurate Description of van Der Waals Complexes by Density Functional Theory Including Empirical Corrections. *J. Comput. Chem.* **2004**, *25* (12), 1463–1473.

(41) Civalleri, B.; Zicovich-Wilson, C. M.; Valenzano, L.; Ugliengo, P. B3LYP Augmented with an Empirical Dispersion Term (B3LYP-D*) as Applied to Molecular Crystals. *CrystEngComm* **2008**, *10* (4), 405–410.

(42) Deringer, V. L.; Stoffel, R. P.; Togo, A.; Eck, B.; Meven, M.; Dronskowski, R. Ab Initio ORTEP Drawings: A Case Study of N-Based Molecular Crystals with Different Chemical Nature. *CrystEngComm* **2014**, *16* (47), 10907–10915.

(43) George, J.; Wang, A.; Deringer, V. L.; Wang, R.; Dronskowski, R.; Englert, U. Anisotropic Displacement Parameters from Dispersion-Corrected DFT Methods and Their Experimental Validation by Temperature-Dependent X-Ray Diffraction. *CrystEngComm* **2015**, *17* (38), 7414–7422.

(44) Hoser, A. A.; Madsen, A. Ø. Dynamic Quantum Crystallography: Lattice-Dynamical Models Refined against Diffraction Data. I. Theory. *Acta Crystallogr. Sect. Found. Adv.* **2016**, *72* (2), 206–214.

(45) Hoser, A. A.; Madsen, A. Ø. Dynamic Quantum Crystallography: Lattice-Dynamical Models Refined against Diffraction Data. II. Applications to l-Alanine, Naphthalene and Xylitol. *Acta Crystallogr. Sect. Found. Adv.* **2017**, *73* (2), 102–114.

(46) Aree, T.; Bürgi, H.-B. Specific Heat of Molecular Crystals from Atomic Mean Square Displacements with the Einstein, Debye, and Nernst–Lindemann Models. *J. Phys. Chem. B* **2006**, *110* (51), 26129–26134.

(47) Wahlberg, N.; Madsen, A. Ø. Implications of X-Ray Thermal Diffuse Scattering in Integrated Bragg Intensities of Silicon and Cubic Boron Nitride. *J. Appl. Crystallogr.* **2017**, *50* (6), 1791–1799.

Atom centered multipolar expansion of the charge density

Piero Macchi
Department of Chemistry and Biochemistry University of Bern, Bern, Switzerland
piero.macchi@dcb.unibe.ch

The conventional modelling of X-ray diffraction data is an atomic expansion of the electron density, namely

$$\rho_{unit\,cell}(\mathbf{r}) = \sum_i \rho_i\,(\mathbf{r} - \mathbf{R}_i)$$

where $\mathbf{R}_i$ is the position of atom *i*. and ρ_i are atomic ground state spherical densities. The superposition of spherical atomic densities represents the so-called *independent atom model* (IAM). $\rho_{unit\text{-}cell}(\mathbf{r})$ is called the *pro-crystal* electron density distribution. If we limit the sum (1) to the atom forming just one molecule in the unit cell, then $\rho(\mathbf{r})$ is the *pro-molecule* electron density.

The atomic multipolar expansion can be regarded as an extension of the spherical atom model and has become the standard for charge density determination in the past 40 years. This is due to the conceptual simplicity and the close relationship with molecular orbital wavefunctions ψ expanded in terms of linear combination of atomic orbitals χ_i (LCAO), that provide the electron density in the form of:

$$\rho(\mathbf{r}) = \sum_\nu \sum_\mu P_{\nu\mu}\chi_\nu(\mathbf{r} - \mathbf{R}_\nu)\chi_\mu(\mathbf{r} - \mathbf{R}_\mu)$$

where

$$P_{\nu\mu} = \sum_j n_j c_{\nu j} c_{j\mu}$$

n_j is the occupancy of the molecular orbital (in a closed-shell system, described by a single configuration, $n_j = 2$ or $n_j = 0$). However, from the LCAO approximation, the one-electron density contains terms of expansion (2), which are centered on *one* atom, as well as *two-center* terms (products of orbitals on two different atoms). The latter are of course more important for the chemical bonding and mostly contribute to the asphericity. An atom-centered multipolar expansion of the electron density is inherently unable to exactly reconstruct the *two-center* terms (Coppens, Willoughby & Csonka, 1971; Coppens, Pautler & Griffin, 1971; Matthews, Stucky & Coppens, 1972; Stewart & Bentley, 1973). However, it is a very practical solution, because it simplifies enormously the model. The o*ne-center* functions in expansion (1) may also be centered at non-nuclear positions, *e.g.* midpoints of the bonds, although an atom centered multipolar expansion is by far the most adopted one. The name *pseudoatom* (Stewart, 1976) was introduced to identify the set of multipole density functions that rigidly move with the atomic nucleus. The functions are parametric as one can optimize the radial behaviour and the

coefficients of an aspherical expansion in terms of spherical harmonics. To refine the parameters, Stewart (1976) proposed the least square fitting of the Fourier transformed electron density, which is directly comparable to the measurable X-ray diffraction structure factors. Gill (1996) called the *pseudoatom* the *Stewart Atom*, which is anyway valid for all the multipolar formalisms proposed over the years, for example by Kurki-Suonio (1977), Hansen & Coppens (1978) and Hirshfeld (1977). By far, the most adopted model is that proposed by Hansen and Coppens, who expanded the atomic density $\rho_i(\mathbf{r})$ as:

$$\rho_i^{H\&C}(\mathbf{r}) = P_{i,core}\rho_{i,core}(\mathbf{r}) + P_{i,valence}\kappa_i^3\rho_{i,valence}(\mathbf{r}) + \sum_{l=0,l_{\max}}\left[\kappa_{i,l}'^3 R_l(\kappa_{i,l}'\mathbf{r})\sum_{m=0,l}P_{lm\pm}y_{lm\pm}(\mathbf{r}/r)\right]$$

This model differs from Stewart's one mainly because of the additional valence monopole, which is missing in Stewart's formalism:

$$\rho_i^{Stewart}(\mathbf{r}) = P_{i,core}\rho_{i,core}(\mathbf{r}) + \sum_{l=0,l_{\max}}\left[R_l(\zeta_l\mathbf{r})\sum_{m=0,l}P_{lm\pm}y_{lm\pm}(\mathbf{r}/r)\right]$$

The difference between the two models is that in (4) the valence is described by two monopoles (typically, one constructed with the radial function derived from a full Hartree-Fock expansion, like for the core density and the other with a single-ζ function, like for the higher multipoles), whereas in (5) the valence monopole as well as the higher multipoles are described with a single-ζ radial density function, see Table 1.

In both equations, the main model parameters (*i.e.* variables of a least square refinement) are κ (or ζ) and P coefficients. The core populations are often kept frozen to nominal values, or collectively refined for all atoms of the same kind in a molecule. While in Hansen & Coppens (1978), the scaling constants κ and κ' are refined, in Stewart (1976), it is the exponent ζ of a Slater function to be directly refined. If the radial function is a single Slater orbital, the two approaches are of course equivalent, whereas they differ if the radial function $R_l(\mathbf{r})$ is a multi-Slater function, as for example adopted to describe d-orbitals in transition metal atoms. The $l+1$ κ_l' parameters are normally constrained to the same value for all poles to avoid divergence. More frequently they are also constrained to be the same for all atoms of a given kind on the structure and it could be sometime necessary to fix $\kappa = \kappa'$, to avoid physically inconsistent values (Abramov *et al.*, 2000).

Angular functions and coordinate system

The choice of angular functions for the multipolar expansion was obviously oriented toward the spherical harmonics. They offer the advantage to be the irreducible representations of the spherical group and form a complete set of orthonormal bases. This has two consequences: a) all parameters of a given atom are linearly independent; b) any product of two spherical harmonics centered on the same atom can be expressed as a linear combination of spherical harmonics (Rose, 1957). In principle, this allows a direct correlation with atomic orbitals, given that the electron density can be expressed as a series of products between orbitals. However, this is not

generally the case, because the spherical harmonics associated with classical multipolar expansions describe with one-center terms also the two-center orbital products. Moreover, orbitals belonging to the same shell (like 2*s*, 2*p* for L shell; or 3*s*, 3*p* and 3*d* for M shell, etc.) are normally associated with very similar radial functions and this makes it impossible to distinguish between poles produced by different orbital products.

The presence of angular oriented functions at an atomic site, implies a definition of axes. The problem is similar to that of the anisotropic displacement parameters in classical refinements of crystal structures. One can of course apply for multipoles the orthogonalized axes of the unit cell as typically adopted for the anisotropic displacement parameters. Instead, Hansen and Coppens (1978) proposed to define a local coordinate system for each atom, which has some advantages: a) it enables to reduce the number of parameters of the multipolar expansion if pseudo-symmetries are exploited (*i.e.* symmetry present in the ideal conformation of an isolated molecule, but not coinciding with symmetry elements of the crystals); b) it enables to refine only multipoles corresponding to a given hybridization state for an atom, see for example Rezende Dos Santos, Genoni & Macchi (2014); c) it enables the exportability of multipoles in appropriate databases, under the hypothesis that atoms of particular functional groups with similar behaviour may behave very similarly. The local system is sometime very intuitive to choose, but often the atomic stereochemistry significantly differs from the ideal one (due to simple hybridization states).

It is important to highlight that a local coordinate system is a flexibility, it is not mandatory and not at all a limiting factor. Analogously, it could be possible defining local coordinate systems also for describing the ADP's, as it is in fact adopted when applying similarity restraints to chemically equivalent atoms or functional groups by standard software for crystal structure refinement.

Radial functions

As provocatively stated by Flensburg, Larsen & Stewart (1995), the choice of a radial function for the atom-centred multipole expansions is "more an art than a science". There are, however, rigid constraints, dictated by physics. For this reason, crystallographers have normally taken radial functions from Slater-type atomic orbitals calculated at Roothan-Hartree-Fock or multiconfigurational level of theory for isolated atoms. In some cases, projected wave functions of numerical solution of four component Dirac-Fock equations have been proposed. This implies analytical functions that mimic those of a Roothan expansion, although not obtained variationally. This proved to be very useful for implementation of relativistic corrections in software that calculates analytical expressions of the electron density, without resorting on more complicated expressions. More recently, calculations using zero order relativistic approximation have been used to obtain multiple-ζ Slater type orbitals for all atoms, see Volkov *et al.* (2006). In Table 7, the most adopted functions, within the multipolar formalism (4) are reported.

Table 1. The most adopted radial functions for multipolar expansion of the electron density, see equations (4) and (5).

Authors	type of wave function	Core	Spherical Valence	Deformation Valence
Clementi & Roetti (1974)	Roothan-Hartree Fock, multiple-ζ atomic orbital functions; non-relativistic	Ideal for elements with low Z	ideal	rarely used
Clementi & Raimondi (1963)	Single-ζ atomic orbital functions, non-relativistic	insufficient	used	ideal
Su & Coppens (1998) and Macchi & Coppens (2001)	Fitting of multiple-ζ Slater atomic orbital functions to Dirac-Fock numerical solution; relativistic	ideal for all atoms	ideal	rarely used
Volkov & Macchi (2006)	multiple-ζ Slater atomic orbital functions, calculated with zero-order regular approximation; relativistic	ideal for all atoms	ideal	rarely used

Physical constraints

The electron density obtained with a multipolar model does not necessarily behave as expected by quantum mechanics. A truly quantum mechanical density is impossible to obtain with that method and the refinement techniques. However, some constraints can be easily applied in order to at least guarantee that some quantum mechanical features are respected:

a) *Electro-neutrality*. This condition is rather obvious and it is obtained by constraining the sum of the valence and core electron populations to be equal to the sum of nuclear changes. Using a least-square procedure, this constraint is quite easily applied. Noteworthy, though, if all core monopole populations and the scale factor are simultaneously refined, the constraint matrix would be singular. For this reason, the scale factor is refined using only the valence electrons as variables or otherwise it is not refined, but calculated from the ratio between the expected number of electrons and those actually obtained, putting as variables also core electrons (Stewart, 1976).

b) *Cusp condition*. According to Kato (1957), the electron density must feature a cusp at the position of each nucleus, satisfying the condition

$$\lim_{r_i \to 0} \left(\frac{\partial}{\partial r_i} + 2Z_i \right) \rho(r) = 0$$

This implies that, in principle, that one should not scale the Slater exponent for orbital *1s*, which is in particular cogent for H atoms. However, this is typically done, applying a κ parameter of 1.2 (or refining it), given that the default exponent ζ = 1.0 (which becomes of course 2.0 in the density representation) produces a too contracted electron density for the H atom. Among the various radial functions, those obtained by non-linear fitting of the numerical solution of a Dirac-Fock equation, like Su & Coppens (1998) and Macchi & Coppens (2001), do not fulfil cusp condition, although they are very close. The deviation from the exact cusp was in fact taken as a criterion to assess the quality of the non-linear fit.

c) *Poisson's condition.* This is dictated by Poisson's electrostatic equation. As shown by Stewart (1977), this implies that the density function must have a radial exponent with $n_l \geq l$ in order to avoid divergence at nuclear positions and therefore violating Poisson's condition. For this reason, the Slater atomic orbital functions must be appropriately modified (if necessary) when used for the radial part of higher poles (deformation density). For example, the orbital radial function of 2s or 2p of a second raw atom, has $n = 1$ which means, in the density formalism, $n_l = 2$. This implies that such a function is inappropriate for an octupole or hexadecapole density function, and it is typically corrected to $n_l = 3$ and 4, respectively. Other recipes have been proposed, especially for third raw atoms and transitional metals, see for example Hansen & Coppens (1978).

d) *Hellman-Feynmann theorem.* This is a well-known and central theorem of quantum mechanics, which implies that the electrostatic forces at nuclei must vanish at equilibrium. Because these forces are mainly due to small dipolar deformations in the vicinity of the nuclei, it is unlikely that with typical resolution of X-ray diffraction measurements one can verify the theorem experimentally. Although in principle feasible, multipolar refinements are typically unconstrained and the standard models (*i.e.* with fixed and spherical atomic cores) return normally small forces at nuclear sites.

f) *Site symmetry*. If an atom sits on a symmetric site, some restrictions apply to the atomic multipoles, following simple rules, as described in Kurki-Suonio (1977). The reason is that, apart for the total symmetric representation of the spherical group (namely, the monopole), all other representation may not be invariant under application of all possible symmetry elements. For example, a dipole is not invariant respect to inversion or to reflection perpendicular to the dipole axis. The application of these rules implies that some multipoles must be necessarily null. Clearly the choice of the local coordinate system must reflect the crystallographic symmetry, otherwise application of the symmetry restriction rules can be very complicated.

g) *Crystal point group symmetry.* Apart from the limitations of the atomic site symmetry, Terpstra, Craven & Stewart (1993) reported on ill determination of odd-poles in non-centrosymmetric space groups. Roversi & Destro (2017) have shown that this depends on a

special class of invariant odd-poles, possible only in non-centrosymmetric crystals, which have phases shifted by $\pm\frac{\pi}{2}$ with respect to the phases generated by the core and valence monopoles (which are the main contributors to the structure factor and those which determine its phase). Using a global unit cell coordinate system, the problem affects, for example, the dipoles $y_{l=1,m_l=0}$ in point groups 3, 4, 6 or the octupoles $y_{l=3,m_l=-2}$ in 222. A solution is that of constraining the sum of all these poles in the asymmetric unit be equal to zero. Noteworthy, the poles mentioned above are defined in the global unit cell coordinate system. If a local coordinate system for each atom is adopted, setting the overall conditions is more difficult. Luckily, for crystals with many atoms in the asymmetric unit (and only few or none on special positions), the effects of the problem of invariant odd-poles are minimized.

Correlation among parameters and uncertainties

To optimize the model parameters, algorithms like those of typical structure refinements are used. At the end of a refinement a variance-covariance matrix is available, which informs us on the uncertainties on each parameter as well as on the correlation among them. The correlation is in fact quite important because it may severely affect the possibility to reconstruct with high precision the electron density of the system. This is due to different reasons:

a) data incompleteness / low resolution: if the number of data is not sufficient (typically at least 10 reflections per parameter are necessary) it is not possible to increase the flexibility of the model (using high value of l_{max}, more radial functions and corresponding κ sets, etc.);
b) if the valence shell of a given atom is too diffuse, only few reflections will be affected by those electrons. This problem is even worse, if the crystal unit cell is small and therefore the sampling of intensities at small values of $|\mathbf{H}|$ is very poor. In this circumstance, a high resolution data set would not be particularly useful to refine the multipolar coefficients of the density functions corresponding to those valence electrons;
c) low precision of the measured data (in this case the weight of some reflections will be systematically lower, reducing *de facto* the number of effective reflections.

While problems a) and c) can be resolved with a more comprehensive and precise data collection, point b) is an inherent problem with the crystal structure and cannot be solved. It mainly affects inorganic compounds, containing few atoms in very small unit cells.

Extended Hansen Coppens models and core refinement

As anticipated above, the radial description of the atomic densities is rather crucial. While atomic wave functions of electronic ground states of isolated atoms provide good approximations, they are not necessarily perfect to describe the accurate electron density distribution. In fact, atomic electron densities contract or expand depending on interactions with other atoms. One can easily predict a shell expansion for atoms bearing a negative charge and contraction for atoms positively charged. Within the multipolar model, this phenomenon is

accounted by modifying the exponents of the Slater Type functions or, in formalism (4), modifying the κ parameters. Much more difficult is taking properly into account the anisotropic deformation due to the chemical bonding. In fact, even within the same electronic shell, one must consider the possibility of different expansion/contractions in different directions. This could be estimated in several ways, but the price would be to significantly increase the number of parameters of the model. A larger flexibility is obtained assigning to each function in the multipolar expansion of equation (4), an independent radial function (Gillet & Koritsanszky, 2012; Koritsanszky, Volkov & Chodkiewicz, 2012):

$$\rho_i(\mathbf{r}) = P_{i,core}\rho_{i,core}(\mathbf{r}) + P_{i,valence}\kappa_i^3\rho_{i,valence}(\mathbf{r}) + \sum_{l=0,l_{\max}}\left[\sum_{m=0,l}\kappa'^3_{i,l,m\pm}R_{l,m\pm}(\kappa'_{l,m\pm}\mathbf{r})P_{l,m\pm}y_{l,m\pm}(\mathbf{r}/r)\right]$$

Equation (7) implies more parameters, because each multipole l,m_l has basically an independent radial function and κ' scaling. In principle, this means $(l_{max}+1)^2$ radial functions (hence κ') parameters for the valence deformation shell, instead of the typically adopted singly κ'. This extra flexibility should be applied with care, because in practical cases instability of the refinement due to large correlation among parameters would be expected. A refinement may be more stable if the m_l-dependent radial functions are rigidly constrained to be the same for each l shell, reducing the number of κ' refined to l_{max} +1.

One could extend the flexibility to the core density as well. As a matter of facts, the possibility to visualize distortions of the atomic cores was proven by Fischer *et al.* (2011). This of course requires a much higher accuracy and resolution compared with standard charge density experiments. The modelling is conceptually simple because one can extend equation (4) to the core electron density. Refining core electron density within a multipolar model is not particularly different from refining the valence electron densities:

$$\rho_i(\mathbf{r}) = \sum_n\left\{P_{i,n}\kappa_{i,n}^3\rho_{i,n}(\mathbf{r}) + \sum_{l=0,l_{\max}}\kappa'^3_{i,n,l}R_{i,n,l}(\kappa'_{i,n,l}\mathbf{r})\left[\sum_{m=0,l}P_{i,n,l,m_l}y_{l,m_l}(\mathbf{r}/r)\right]\right\}$$

where in principle the extra radial flexibility of equation (7) could also be adopted.

Anyway, a strong correlation problem may be easily envisaged even if the radial flexibility is limited to l-dependency. As summarized by Macchi (2013), the core charge density refinement could be carried out with different degrees of flexibility, namely:

(1) *minimal*: refining the atomic core monopole populations in a typical multipolar refinement, that means making P_{core} in equation (4) a variable;

(2) *semi-flexible*: refining a scale κ_{core} factor together with the monopole population, allowing therefore a contraction/expansion of the core itself;

(3) *flexible*: refining a full set of multipoles (even up to hexadecapole) for the core electrons, starting from the orbitals of the atomic wave functions.;

(4) *extremely flexible*: refining different sets of multipoles and contraction factors for each electronic shell of the core for atoms of the third period or higher (thus one set of multipoles for K shell, one for L-shell etc.)..

Databases approaches

In contrast with the extra-flexibility of the extended Hansen-Coppens models, one could instead reduce the flexibility of (4) when dealing with lower quality data and very large systems. In this respect, the local definition of a coordinate system is very useful to adopt the concept of atomic electron density *transferability*. The assumption is that the stereochemistry of an atom (hybridization state, number and type of bonds made by the atom, conformation of the functional group to which the atom belongs) mostly determines the electron density distribution of the atom, hence the derived multipolar expansion. This implies that the same set of atomic multipoles can equally well describe a given functional group inserted in different molecules, and packed in different crystals.

The purposes of this kind of studies are: a) improve the structural refinement of a molecule even when the measured data are of quality insufficient for a full charge density study, or when the resolution is not sufficient to include so many parameters; b) having a rapid evaluation of electrostatic properties of molecules even when fully free multipolar refinements cannot be carried out.

Over the years, several databases have been developed (Pichon Pesme, Lecomte & Lachekar, 1995; Volkov *et al.*, 2004; Dittrich *et al.*, 2006), based on different grounding but having in common the same idea of transferability. Pichon Pesme, Lecomte & Lachekar (1995) built the database of multipoles from a series of experimental studies on amino acids and peptides or other molecules containing same functional groups. A similar idea was developed by Volkov *et al.* (2004), however using multipoles calculated with theoretical methods and expanded in terms of atomic multipoles. Instead, Dittrich *et al.* (2006) proposed the so-called invariant pseudoatom or *invariom* (invariant with respect to the transfer from a general model to an actual molecule). An atom type is identified from a generalized definition: the nearest neighbours of the *invariom* are the same as for the atom in the actual molecule. Thus, a theoretical calculation on a prototype molecule (constructed with the *invariom* principle) enables defining the *invariom* set of multipoles for an atom.

References

Abramov, Y. A., Volkov, A., Wu, G. & Coppens, P. (2000). *J. Phys. Chem. B*, **104**, 2183-2188.

Coppens, P., Pautler, D. & Griffin, J. F. (1971). *J. Am. Chem. Soc.*, **93**, 1051-1058.

Coppens, P., Willoughby, T. V. & Csonka, L. N. (1971). *Acta Cryst.*, **A27**, 248-256.

Clementi, E. & Raimondi, D. L. (1963). *J. Chem. Phys.*, **38**, 2686-2689.

Clementi, E. & Roetti, C. (1974). *Atomic Data and Nuclear Data Tables*, **14**, 177-478.

Dittrich, B, Hubschle, C. B., Luger, P & Spackman, M.A. (2006).. *Acta Cryst.*, **A62**, 1325-1335.

Dos Santos, L.; Genoni, A. & Macchi, P. (2014). *Acta Cryst.*, **A70**, 532-551.

Fischer, A., Tiana, D., Scherer, W., Batke, K., Eickerling, G., Svendsen, H., Bindus, N. & Iversen, B. B. (2011). *J. Phys. Chem. A.* **115**, 13061–13071.

Flensburg, C., Larsen, S. & stewart, R. F. (1995). *J. Phys. Chem.*, **99**, 10130-10141.

Gill, P. M. W. (1996). *J. Phys. Chem.* **100**, 15421-15427.

Gillet, J. M. & Koritsanszky, T. (2012). *Past, Present and Future of Charge Density and Density Matrix Refinements.* in *Modern charge density analysis.* Gatti, C., Macchi, P. editors. Springer.

Hansen, N. K & Coppens, P. (1978). *Acta Cryst.* **A34**, 909–921.

Hirshfeld, F. L. (1977a). *Isr. J. Chem.* **16**, 198–201.

Hirshfeld, F. L. (1977b). *Isr. J. Chem.* **16**, 226–229.

Hirshfeld, F. L. (1977c). *Theor. Chim. Acta.* **44**, 129–138.

Koritsanszky, T., Volkov, A. & Chodkiewicz, M. (2012). *New Directions in Pseudoatom-Based X-Ray Charge Density Analysis.* in *Electron Density and Chemical Bonding,* Ed. Stalke, D., special issue of *Structure and Bonding*, **146**, 1-26.

Kurki-Suonio, K. (1977a). *Isr. J. Chem.* **16**, 115–123.

Kurki-Suonio, K. (1977b). *Isr. J. Chem.* **16**, 132–136.

Macchi, P. (2013). *Cryst. Rev.*, ***19***, 58-109.

Macchi, P. & Coppens, P. (2001). *Acta Cryst.*, **A57**, 656-662.

Matthews, D. A., Stucky, G. D. & Coppens, P. (1972). *J. Am. Chem. Soc.*, **94**, 8001-8008.

Pichon Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.*, **99**, 6242-6250

Rose, M. E. (1957). *Elementary Theory of Angular Momentum.* New York: Wiley.

Roversi, P. & Destro, R. (2004). *Chem. Phys. Lett.*, **386**, 472-478.

Stewart, R. F. (1976). *Acta Cryst.* **A32**, 565–574.

Stewart, R. F. (1977). *Isr. J. Chem.*, **16**, 124-131.

Stewart, R. F., & Bentley (1973). *J. Comput. Phys.*, **11**, 127-145.

Su, Z. & Coppens, P. (1998). *Acta Crys.*, **A54**, 646-652.

Terpstra, M., Craven, B. M. & Stewart, R.F. (1993). *Acta Cryst.*, **A49**, 685-692.

Volkov, A., Koritsanszky, T. & Coppens, P. (2004). *Chem. Phys. Lett.*, **391**, 170-175.

Volkov, A. & Macchi, P. (2006) *Zero-order regular approximation atomic wavefunctions for multipolar refinement*. in Volkov, A., Macchi, P., Farrugia, L. J., Gatti, C., Mallinson, P., Richter, T. & Koritsanszky, T. (2006). *XD2006 - A Computer Program Package for Multipole Refinement, Topological Analysis of Charge Densities and Evaluation of Intermolecular Energies from Experimental and Theoretical Structure Factors.*

Neutron diffraction and spin density multipolar model

Nicolas Claiser
Université de Lorraine, CRM2 (UMR UL-CNRS 7036), Vandoeuvre-lès-Nancy, France
nicolas.claiser@univ-lorraine.fr

Maxime Deutsch
CRM2 Laboratory, Lorraine University, Vandoeuvre-lès-Nancy, France
maxime.deutsch@univ-lorraine.fr

Neutron Diffraction

UnPolarized Neutron Diffraction

Properties of the neutron

Mass: 1.675×10−24 g

Charge: 0

Life time : 886.8 s (14.8 minutes)

Spin:1/2

Interaction with crystals

The scattering length of the neutron-nucleus system is the basic quantity which describes the strength and character of the interaction of neutrons with the individual nuclei. The values of scattering lengths vary irregularly from one nucleus to another in the periodic classification. Therefore neutrons are an important tool for the investigation of the static and dynamic properties of condensed matter since they distinguish between various elements and isotopes.

Neutron scattering lengths

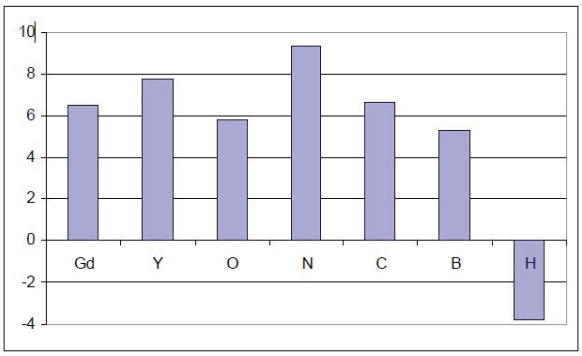


Figure 1. Nuclear scattering lengths (10^{-15}m) of some atoms.

Comparison between X-rays and neutrons

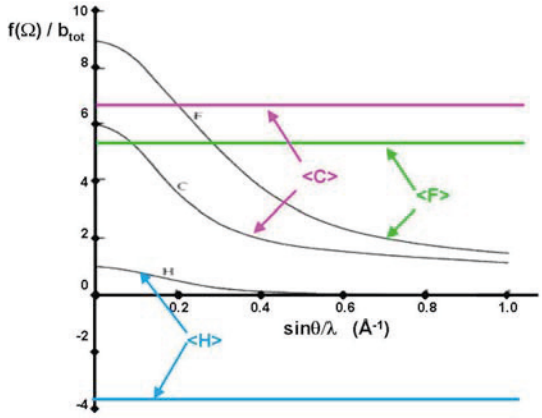


Figure 2. Comparison of nuclear diffusion factor and nuclear scattering length (10^{-15}m) of some atoms.

Polarization neutron diffraction

Polarization of a neutron beam and action of a magnetic field

The spin polarization of the neutron beam is a classical vector. In a magnetic field, the magnetic moment of the neutron precesses around the field. If the direction of this field changes abruptly, the polarization vector replaces it rotation around this new orientation.

Interaction with crystals

For a beam of polarised neutrons, the differential scattering cross sections (σ) corresponding to a Bragg reflection can be expressed as functions of the nuclear and magnetic structure factors (F_N and F_M, respectively, **Q** being the diffraction vector):

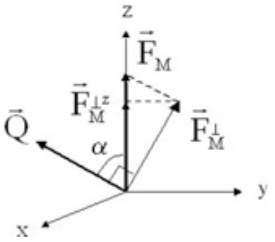
$$\left(\frac{d\sigma}{d\Omega}\right)^{\pm\pm} \propto |F_N \pm F_M^{\perp z}|^2 \qquad F_N(\mathbf{Q}) = \sum_i^{atoms} b_i \, e^{i\mathbf{Q}\cdot\mathbf{r}_i} \, e^{-W_i}$$

$$\left(\frac{d\sigma}{d\Omega}\right)^{\pm\mp} \propto |F_M^{\perp x} \pm i F_M^{\perp y}|^2 \qquad \mathbf{F_M}(\mathbf{Q}) = \int_{unit\ cell} \mathbf{M}(\mathbf{r}) \, e^{i\mathbf{Q}.\mathbf{r}_i} \, d\mathbf{r}$$

Without any polarization analysis of the diffracted beam the measured intensity is:

$$I(\mathbf{Q})_{\pm} \propto \left[F_N^2 \pm (F_N^{\star} F_M^{\perp z} + F_N F_M^{\perp z^{\star}}) + F_M^{\perp\,2}\right]$$

$F_M^{\perp}$ is a vector corresponding to the component of the magnetic structure factor (F_M) perpendicular to the diffraction vector. $F_M^{\perp z}$ is the vertical component of the $F_M^{\perp}$ vector (i.e. parallel to the magnetic field applied to the sample) : $F_M^{\perp z} = \sin\alpha \, F_M^{\perp} = \sin^2\alpha \, F_M^{\perp}$



The polarized neutron diffraction (PND) experiment

This experiment requires a neutron source (nuclear reactor, spallation source) and devoted beam-lines allowed to maintain the polarization of the incident beam. The sample itself is subject to an external magnetic field. A schematic description of the 5C1 beamline at CEA Saclay (France) is given by figure 3.

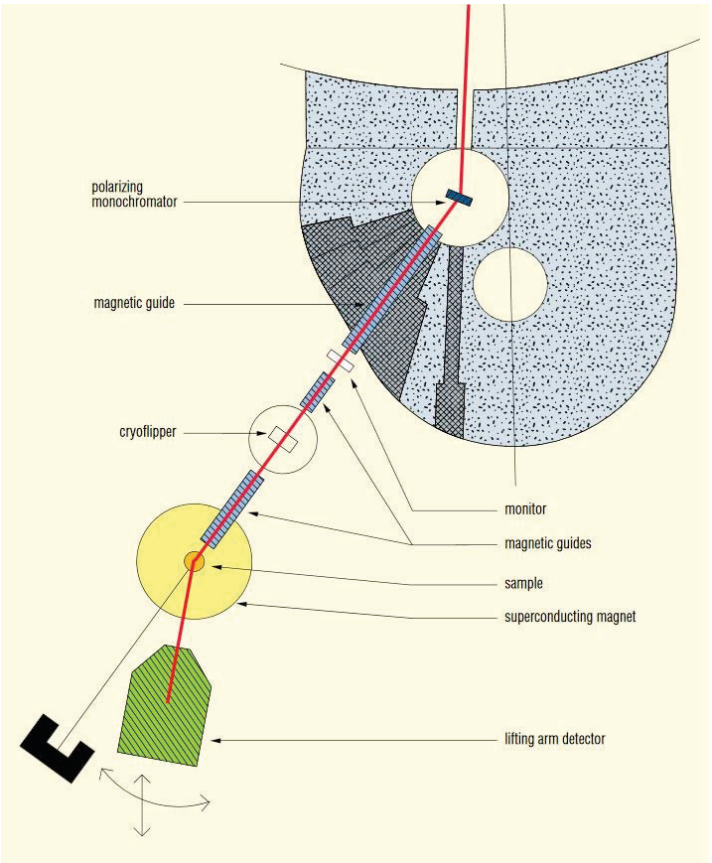


Figure 3. Schematic representation of the 5C1 beamline at LLB (CEA Saclay).

The polarized neutron diffraction (PND) technique

The PND technique takes advantage of the incident polarization dependency of the cross-section to measure precise quantitative magnetization distributions of single crystals. This

technique is mainly used for investigating single crystals that are ferro- or ferri-magnetically ordered in an applied magnetic field. It can also be applied to some antiferromagnetic materials. The magnetic field is vertically applied to the studied crystal. The beam is supposed to be perfectly polarized parallel (+) or anti parallel (-) to this applied field. Then, for each reflection of a centrosymmetric crystal, we measure the flipping ratio R(**Q**) which is define as the ratio between the cross-sections measured for polarization (+) and (-):

$$R(\mathbf{Q}) = \frac{I_+}{I_-}$$

The data treatment and correction

- Correction due to the beam: a part of the corrections are due to the imperfect polarization of the incident beam (P) or to the efficiency of the flipping. The polarization correction is linked to the part of the signal (Φ) mesured for each polarisation:

$$P = \frac{\phi^{\uparrow} - \phi^{\downarrow}}{\phi^{\uparrow} + \phi^{\downarrow}}$$

- Correction due to the sample: extinction is different for I_+ and I_-

- Correction due to the nuclear polarization: Two effects have to be considered. First, the interaction between the spin of the neutron and the nuclear spin which is proportional to the polarisation of the nuclear spins. This effect is particularly sensitive for hydrogen atoms. The second correction is due to the interaction between the magnetic moments of the neutrons and the electric field created by the electrons and nucleus (Schwinger effect)

Spin density modeling

There are two possible ways to recover the spin density from the experimental data:

- directly without making assumptions about the nature of this distribution
- with a parametric model

Direct Methods

Fourier transform

The spin density can be written as the inverse Fourier transform of experimental magnetic structure factors provided that the sum is infinite. In practice, this sum is limited by the experimental limit of $\sin(\theta/\lambda)$. Furthermore, only Bragg reflections for which the nuclear structure factor F_N is sufficient can be measured. Therefore, this method is not convenient to obtain reliable density.

Maximum entropy methods

The Maximum Entropy Method (MEM) consist of maximizing entropy and at the same time fulfills the condition $\chi^2 = 1$. [Papoular1990]

To calculate the entropy, the lattice of the compound is divided in M pixels in which a constant density is assumed. For the spin we consider a double distribution of positive quantities n_i =

$\rho^+(\mathbf{ri})$ and $n_{i+M} = \rho^-(\mathbf{ri})$, the magnetization density is then given by $s_i = \rho^+(\mathbf{ri})-\rho^-(\mathbf{ri})$. The entropy of the density in the magnetization is then defined by:

$$S(s) = -\sum_{i=1}^{2M} p_i Log(p_i) \quad \text{with} \quad p_i = \frac{\rho_i}{\sum_{j=1}^{2M} \rho_j}$$

and χ^2 is defined by:

$$\chi^2(s) = \frac{1}{N}\sum_{Q} \frac{|F_M^{exp}(Q) - F_M^{calc,s}(Q)|^2}{\sigma^2(Q)}$$

Parametric model

Different approaches can be considered depending on the quantity being modeled: the unpaired electron wave function or the spin density [Schweizer2001]. In both cases, the model consists of an analytic expression, with parameters that are refined by comparing theoretical magnetic structure factors (or flipping ratios in non-centro cases) with experimental values. These parameters are then used to reconstruct the spin density with the analytical model.

Wave function Model

In the Restricted Hartree-Fock (RHF) description of the wave function of a molecular system with an unpaired electron, the spin density is given by the square of the modulus of the molecular orbital $\phi(\mathbf{r})$ occupied by a single electron:

$$s(\boldsymbol{r}) = |\phi(\boldsymbol{r})|^2$$

Where the molecular orbital ϕ φ is written as a linear combination of atom centered atomic orbitals $\psi_i(r_i)$:

$$\phi(\boldsymbol{r}) = \sum_{i}^{atoms} \alpha_i \psi_i(\boldsymbol{r})$$

Atomic orbitals are themselves linear combinations of Slater type atomic functions.

In the final expression of the magnetic structure factors (not shown here), the parameters to be refined are the atomic spin populations together with the atomic orbital coefficients.

Multipolar Model

Multipolar modeling of spin density has been derived from the charge density Hansen and Coppens formalism []. The spin density is described as a sum of atomic densities:

$$s(\boldsymbol{r}) = \sum_{i} s_i^{mult}(\boldsymbol{r}_i)$$

And the atomic spin density is modelled thanks to a multipolar functions basis:

$$s_i^{mult}(\boldsymbol{r}_i) = P_v^i \kappa^3 R_0^i(\kappa r_i) + \sum_{l=0}^{lmax} \kappa'^3 R_l^i(\kappa' r_i) \sum_{m=0}^{l} P_{lm\pm}^i y_{lm\pm}(\theta,\varphi)$$

With $R_l^i(\kappa r)$ a Slater type function (κ being a contraction coefficient that can be refined) and y_{lm} are real spherical harmonics.

This expression is similar to the Hansen & Coppens charge density model without any core contribution.

Joint Refinement

Joint model: Extended Hansen & Coppens model [Deutsch2012]

As previously seen charge and spin density were modeled thanks to similar multipolar models. Therefore, a joint refinement of the spin and charge densities in an extended model which distinguishes the up and down spin contributions, is possible. In this new model two types of atoms have to be distinguished: non-magnetic atoms, which are refined with a classical Hansen & Coppens model and magnetic atoms, for which all the density parameters are split. For magnetic atoms, the valence (P_v), multipolar ($P_{lm\pm}$) populations and expansion/contraction parameters (κ) should be split in up and down counterparts. The extended model is described as:

$$\rho(\mathbf{r}) = \rho_{core}(r) + P_v^{\uparrow}\,\kappa^{\uparrow 3}\,\rho_v^{\uparrow}(\kappa^{\uparrow}\mathrm{r}) + P_v^{\downarrow}\,\kappa^{\downarrow 3}\,\rho_v^{\downarrow}(\kappa^{\downarrow}\mathrm{r}) \quad + \sum_{l=0}^{l_{max}} \kappa'^{\uparrow 3}\,R_l(\kappa'^{\uparrow} r) \sum_{m=0}^{l} P^{\uparrow}{}_{lm\pm}\, y_{lm\pm}(\theta,\phi) + \sum_{l=0}^{l_{max}} \kappa'^{\downarrow 3}\,R_l(\kappa'^{\downarrow} r) \sum_{m=0}^{l} P^{\downarrow}{}_{lm\pm}\, y_{lm\pm}(\theta,\phi)$$

where $\uparrow$ and $\downarrow$ state for spin up and down parameters. $\kappa^{\uparrow}/\kappa^{\downarrow}$ were introduced because the spin up and spin down electron distributions may not have the same radial extension [Becker1985]. This joint refinement against XRD and PND data leads to a simultaneous determination of spin and charge density distributions with higher level of details for spin distribution thanks to XRD data constraints.

The spin density is obtained by calculating the difference between spin up and spin down densities:

$$s(\mathbf{r}) = P_v^{\uparrow}\,\kappa^{\uparrow 3}\,\rho_v^{\uparrow}(\kappa^{\uparrow}\mathrm{r}) - P_v^{\downarrow}\,\kappa^{\downarrow 3}\,\rho_v^{\downarrow}(\kappa^{\downarrow}\mathrm{r}) \quad + \sum_{l=0}^{l_{max}} \kappa'^{\uparrow 3}\,R_l(\kappa'^{\uparrow} r) \sum_{m=0}^{l} P^{\uparrow}{}_{lm\pm}\, y_{lm\pm}(\theta,\phi) - \sum_{l=0}^{l_{max}} \kappa'^{\downarrow 3}\,R_l(\kappa'^{\downarrow} r) \sum_{m=0}^{l} P^{\downarrow}{}_{lm\pm}\, y_{lm\pm}(\theta,\phi)$$

This Extended Hansen & Coppens model was implemented in MOLLYNX software.

Weighing of different data sets

One of the most intriguing question about the joint refinement is how to weight such different experiment with different statistics and uncertainties? Indeed, large difference between the numbers of reflections for each experimental data set is generally observed (~10,000 for XRD and ~100 for PND) [Deutsch2012, Deutsch2013].

Actually three weighting schemes were envisioned in order to manage these differences:

- **UNIT**, where the score function C minimizes the sum of the χ^2 of each experiment; this model was used in the joint refinement (UND and XRD) by Coppens and co-workers [Coppens1981]:

$$C(\chi_j^2(x)) = \sum_j \chi_j^2(x)$$

where j stands for an experiment (XRD, UND, or PND) and

$$\chi_j^2(x) = \sum_i \left| F_i^{jO} - F_i^{jC}(x) \right|^2 / \sigma^2(F_i^{jo})$$

where i runs over all the measured structure factors Fo; Fc are the calculated ones and σ^2 is the estimated variances of Fo.

- **NLOG** for which the score function C is defined as:

$$C(\chi_j^2(x)) = \sum_j N_j \log(\chi_j^2(x))$$

where Nj is the number of observation of data set j

This scheme was proposed by Bell *et al* [Bell1996] and Gillet et al [Gillet2004] based on the logarithm of χ^2 to reduce the weighting ratio between large and small data sets and hence to better take into account difference in uncertainties evaluation .

The third weighting scheme was proposed to favors even more the small data set by giving approximately the same weight for the small and the big data sets

- **LOG**: a new weighting scheme proposed, independent from the data set size, where the score function is:

$$C(\chi_j^2(x)) = \sum_j \log(\chi_j^2(x))$$

Because of the large difference between the numbers of reflections for each experiment the NLOG or LOG scheme may prevent from neglecting the experiments with small size data collections (i.e. PND data set).

The initial model for the joint refinement is the density model obtained by X-ray multipolar refinement only. For atoms supposed to carry a spin density, their valence and multipole populations are then split into up and down and refined against all data sets.

Assumptions and constraints

Several assumptions and constraints have to be done in order to obtain physically meaningful results:

- The cell parameters are those obtained from the X-ray experiments (generally more precise due to the larger number of measured reflections).
- Two sets of anisotropic atomic displacement parameters Uij and extinction parameters are refined from X-ray and neutron separately due to the possible difference of crystal size or difference in the temperatures of the data collections.
- The electroneutrality (all X-rays monopoles) and number of unpaired electrons (spin monopoles) constraints are added.

Conclusion: example of joint refinement application Cu2

A joint refinement of X-ray neutron and polarized neutron diffraction data were done on a di-azido cooper complex [Deutsch2014]

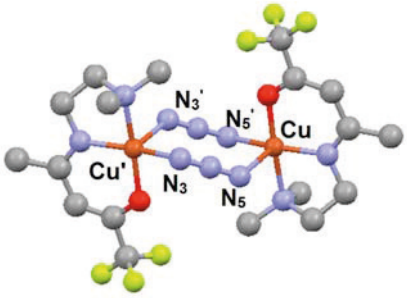


Figure 4. View of the di-azido copper complex. N atoms are represented in blue, O in red, C in grey, F in yellow and Cu in orange. H atoms are not shown for sake of clarity.

The major results of this joint refinement are presented in figure 5 and figure 6 below.

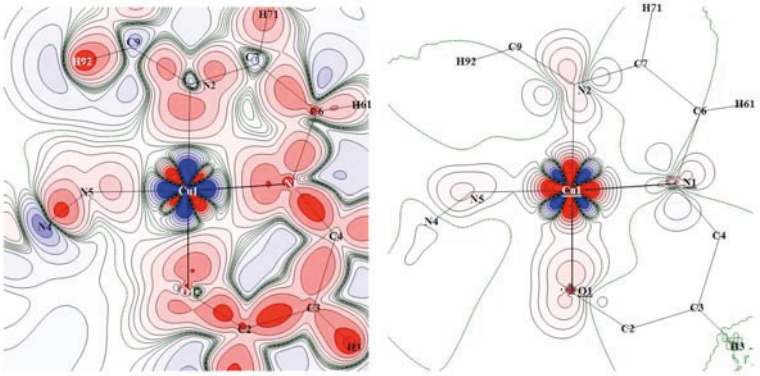


Figure 5. Charge and spin density maps in the plane containing Cu, O1 and N5. (a) Static deformation density map obtained by means of the joint refinement strategy. Isocontours are drawn for $0.01*2^n$ e.Å^{-3} with n =0–13 (positive red, negative blue). (b) Spin density map obtained by means of the joint refinement strategy. Isocontours are drawn for $0.01*2^n$ μB.Å^{-3} with n = 0–13, spin up contours in red, spin down contours in blue.

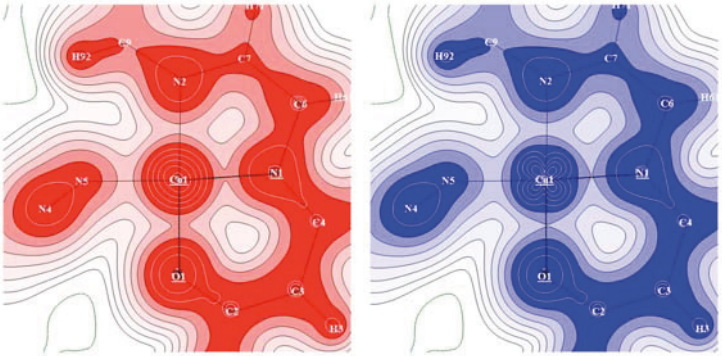


Figure 6. Spin-resolved electron densities. Left: (a) Experimental spin up (majority) and (b) experimental spin down (minority) valence electron densities from joint refinement of the spin-split model. The density distributions are represented in the Cu—N1—O1 plane (contours $0.01*\ 2^n$ e.Å^{-3} (n = 0–12)).

It clearly appears that the Extended Hansen & Coppens model, is successful at precisely retrieving all the essential features of the electron distribution. Most important is the dramatic difference between spin up and spin down angular distributions in the vicinity of the copper nuclei (see figure 6) in perfect agreement with theoretical calculations.

References and Bibliography

- Neutron data booklet second edition https://www.ill.eu/fileadmin/user_upload/ILL/1_About_ILL/Documentation/NeutronDataBooklet.pdf
- Becker, P. & Coppens, P. (1985). Acta Cryst. A 41, 177–182.
- Bell, B., Burke, J., & Schumitzky, A. (1996). Comp. Stat. And Data Anal 22, 119–135.
- Brown P.J. , J. B. Forsyth and R. Mason, Phil. Trans. R. Soc. Lond. B, vol. 290, pp. 481-495, 1980.
- Coppens, P., Boehme, R., Price, P. F., & Stevens, E. D. (1981). Acta Cryst. A37, 857–863.
- Deutsch M. et al, Acta Cryst. Section A, vol. 68, no. 6, pp. 675-686, 2012.
- Deutsch M. et al IUCrJ. - 2014. - Vol. 1. - pp. 194-199.
- Deutsch M., Claiser N., Souhassou M. & Gillon B. (2013) Physics Procedia 42, 10-17
- Gillet, J.-M. & Becker, P. J. (2004). J. Phys. Chem. Solids, 65, 2017 - 2023.
- Hansen, N. K. & Coppens, P. (1978). Acta Cryst. A34, 909–921.
- Papoular, R. and Gillon, B., Europhys. Lett. 13 (1990) 429.
- Schweizer, J., Ressouche, E., in 'MagnetoScience From Molecules to Materials', Eds Miller, J., Drillon, M., Wiley, (2001) pp. 325–355.

For French reader

1. Maxime Deutsch phd Thesis: https://hal.archives-ouvertes.fr/tel-00924396
2. B. Gillon "La technique classique du rapport de flipping. Application aux aimants moléculaires et aux aimants photo-commutables" Collection SFN 7 (2007) 13–40 EDP Sciences, Les Ulis DOI: 10.1051/sfn:2007017

Experimental Quantum Chemistry

Martin Rahm
Chalmers University of Technology, Department of Chemistry and Chemical Engineering, Gothenburg, Sweden
martin.rahm@chalmers.se

Introduction

What are the underlying reasons for a chemical reaction? Understanding the connection between electronic structure and material properties in a way that allows for "chemical thinking" and rational design is a principal challenge of chemical science. This contribution describes a conceptual framework called "Experimental Quantum Chemistry" that aims to bridge the gap between quantum chemical calculations and experimental measurements, including photoelectron and vibrational spectroscopy, X-ray structure determination, electron density mapping, as well as measurements of energies of activation, reaction and formation.
Before going into details and examples, a brief summary to complementary approaches is warranted. In addition to older concepts, such as Lewis structures, electronegativity, acidity scales, and atomic radii, there exist, in principle, two overreaching approaches for analyzing electronic structure:

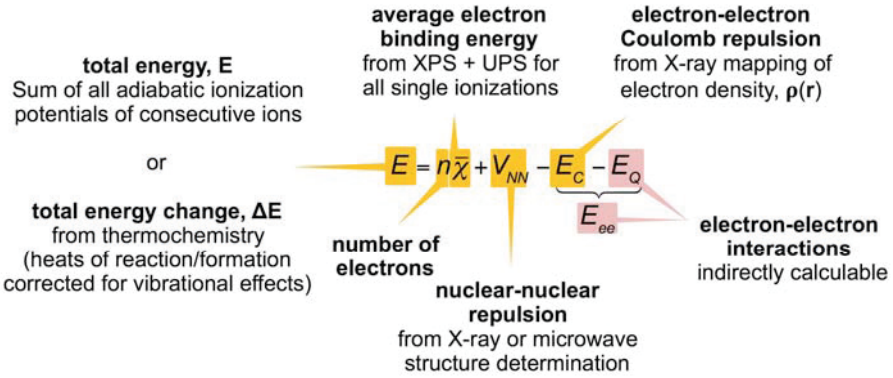
Wavefunction-based Analyses: Molecular orbital theory arose from simplified physical models, such as Hückel theory,[1–5] and is the premier method for calculating, analyzing, conceptualizing and predicting chemical transformations to date. Modern valence bond theory is very much complementary to this.[6] Together with orbital localization procedures, these quantum chemical frameworks allow for different delocalized and localized bonding schemes.[7–9] Valuable crossover methodology between molecular and extended systems, such as polymers and crystals, exists, including orbital projection methods, allowing for spatially localized concepts, such as atomic charge, orbital hybridization and bonding-antibonding character.[10]
Energy decomposition analyses (or EDAs),[11–17] is another set of elegant methods that can provide detailed insight into chemical bonding. One goal of EDA methods is often to provide definitions for as many interpretable energy contributions as possible (Pauli repulsion, dispersion, electrostatics, orbital relaxation, etc.).[18] As a consequence, EDAs typically requiring iterative processes, projections schemes or orbital localization and rotational procedures.
A large number of chemical descriptors, such as electronegativity, hardness, softness and different reactivity indices, have also been defined within the framework of conceptual Density Functional Theory, which quantifies responses in energy with respect to different perturbations.[19, 20] All of these approaches and methods are useful for analyzing chemistry, but they share one thing in common – they all require at least one quantum mechanical calculation to approximate a wavefunction, or a density.

Quantum Chemical Topology (QCT): Several methodologies have been developed for analyzing the topology of densities in three dimensions, which, depending on their type, can be a consequence of the underlying electronic structure. The most prominent approach is the Quantum Theory of Atoms in Molecules (QTAIM),[21] which established a paradigm for topological analysis of electron densities. Because the electron density is both experimentally observable and quantum mechanically calculable,[22] QTAIM is an inherently interdisciplinary approach. "Interacting Quantum Atoms" is one EDA method defined for the QTAIM framework.[13, 23] Today, Quantum Chemical Topology is a rapidly growing field of research that includes the study of several kinds of densities,[24–26] representing, for example, aspects of electron localization,[27–30] and non-covalent interactions.[31, 32] Quantum Crystallography, the topic of this summer school, is a related field of research, in which X-ray data is used to constrain calculations of electronic wavefunctions.[33, 34]

Whereas there certainly are challenges in the experimental QCT field,[35,36] it is becoming increasingly clear that QCT methods in general, along with orbital analyses, and intricate EDA schemes are all highly complementary.[37] Today, these different approaches are often used in tandem in chemistry and materials science.[38,39]

Introduction to Experimental Quantum Chemistry

At the core of "Experimental Quantum Chemistry" (EQC) is an energy partitioning that reads as:[40]



Eq. 1, here omitting thermal and entropic effects, is, in principle, general and exact within the Born-Oppenheimer approximation. As will be shown, this partitioning allows interchangeable use of both theory and experimental data in a single framework to describe energy processes.[40] Practically, each term of Eq 1 can be *approximated* using experiments as indicated, or be calculated using quantum mechanical methods, including any Wave Function Theory or Density Functional Theory (DFT) methods.[40,41] Both approaches will be described. Following this, interpretations of the different terms will be discussed alongside some examples. Finally, a chemical bonding descriptor will be introduced.

Computational Approach

To understand why Eq. 1 holds it helps to express its terms at different levels of quantum mechanical approximation. As a useful example, we first write the equivalent partitioning to Eq. 1 within Hartree-Fock theory. The Hartree-Fock energy expression can be formulated such that,

$$E_{(HF)} = n\bar{\chi}_{(HF)} + V_{NN} - E_{ee(HF)}, \tag{2}$$

which compares to Eq. 1. The first term, the average electron binding energy, $\bar{X}$, is then expressed as:

$$\bar{\chi}_{(HF)} = n^{-1}\sum_{i=1}^{n}\left(\left\langle \phi_i \middle| \hat{h} \middle| \phi_i \right\rangle + \sum_j (J_{ij} - K_{ij})\right) = n^{-1}\sum_{i=1}^{n}\varepsilon_i \tag{3}$$

where ϕ_i stands for occupied spin orbitals, and $\hat{h}$ is the one electron operator of the hamiltonian J_{ij} and K_{ij} are the matrix elements of the Coulomb and exchange operators. This approximation to $\bar{X}$ invokes Koopmans' theorem and equals the average of the eigenvalues, ε_i, of all occupied molecular orbitals. One general advantage with this approach is that $\Delta\bar{\chi}$, i.e. the change in the average electron binding energy over the course of a transformation, appears rather insensitive to the level of theory.[40] The interpretation of $\bar{\chi}$ and $\Delta\bar{\chi}$ will be discussed in more detail below. The second term of Eq. 2, the nuclear-nuclear repulsion, is a direct consequence of the molecular structure, and calculates as the classical Coulomb repulsion term:

$$V_{NN} = \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}, \tag{4}$$

where M is the number of nuclei. The third term of Eq. 2, the average electron-electron interaction energy, E_{ee}, writes as:

$$E_{ee(HF)} = \frac{1}{2}\sum_j (J_{ij} - K_{ij}), \tag{5}$$

which shows that E_{ee} only represents electron-electron interactions, expressed in terms of the same Coulomb and Exchange operators that enter $\bar{\chi}_{(HF)}$. Note that just like in Eq. 1, the $E_{ee(HF)}$-term shows up with a negative sign in Eq 2. This is to correct for the double counting of electron-electron repulsion in the $\bar{\chi}$-term, and is a general consequence of formulating the energy expression in this manner.

Estimates to the total energy E (or ΔE of a transformation) can, of course, be sensitive to how accurately correlation effects are treated. Such effects are typically better captured using DFT. The corresponding EQC-partitioning of the KS-DFT energy expression reads as:

$$E_{(KS)} = n\bar{\chi}_{(KS)} + V_{NN} - \underbrace{\left(V_{ee}(\rho) + \int \frac{\delta E_{XC}(\rho)}{\delta\rho(r)}\rho(r)dr - E_{XC}(\rho)\right)}_{E_{ee(KS)}}, \quad (6)$$

where $V_{ee}(\rho)$ is the classical electron-electron Coulomb repulsion energy, E_{XC} is the exchange-correlation energy, and the middle term in the $E_{ee(KS)}$-expression is the exchange-correlation potential. For extended structures, $\bar{\chi}$ can be obtained from the density of states (DOS) as,

$$\bar{\chi} = \frac{\int_{-\infty}^{\varepsilon_f} E \times DOS(E)dE}{\int_{-\infty}^{\varepsilon_f} DOS(E)dE}, \quad (7)$$

where ε_f is the Fermi energy.

It is also possible to define $\bar{\chi}$ more generally without the explicit need for orbitals,

$$\bar{\chi}_{gen} = n^{-1}\int\left(\tau_L(\mathbf{r}) + \upsilon(\mathbf{r})\rho(\mathbf{r}) + 2\int\frac{P(\mathbf{r},\mathbf{r}_2)}{|\mathbf{r}-\mathbf{r}_2|}d\mathbf{r}_2\right)d\mathbf{r}, \quad (8)$$

where $\tau_L(\mathbf{r})$ is the Laplacian form of the kinetic-energy density, $\nu(\mathbf{r})$ is the external potential, $\rho(\mathbf{r})$ is the electron density, and $P(\mathbf{r},\mathbf{r}_2)$ is the diagonal of the two-electron reduced density matrix, all which can be extracted from any single- or multi-reference wavefunction. In multi-reference descriptions, the interpretation of $\bar{\chi}$ in terms of ionization potentials becomes approximate, but its interpretation as an inherent average of electron binding energies remains. Within this coordinate representation, the analogous expression to Eq. 1 reads as:

$$E_{gen} = n\bar{\chi}_{gen} + V_{NN} - \underbrace{\int\frac{P(\mathbf{r},\mathbf{r}_2)}{|\mathbf{r}-\mathbf{r}_2|}d\mathbf{r}_2}_{E_{ee}} \quad (9)$$

Experimental Approach

All of the different terms of Eq. 1 can be estimated experimentally. Total energies, E, can be obtained as a summation of adiabatic ionization potentials of an atom or molecule.[40] For example, for helium, E = IP(He) + IP(He^+).

Whereas E can be obtained experimentally for some systems, it is often more relevant to consider relative energies. ΔE determine the outcomes of most chemical and physical processes, and can correspond to a chemical reaction, such as a bond formation, the thermodynamics of a combustion reaction, the kinetics of passing over a transition state barrier, or the vibration about a bond equilibrium. For chemical transformations, estimates of ΔE requires accurate

thermochemical data (for example, heats of formation). These can be corrected for vibrational effects by identifying fundamental vibrational frequencies, ν_i, from IR and Raman spectroscopy,

$$\Delta E \approx \Delta H^0 - \Delta E_{ZPE}\,, \tag{10}$$

where E_{ZPE} can be estimated within the harmonic approximation as $\frac{1}{2}h\sum\nu_i$.[42]
Experimentally, the average electron binding energy, $\bar{\chi}$, can be estimated as:

$$\bar{\chi} = n^{-1}\sum_{i=1}^{n} d_i\varepsilon_i\,, \tag{11}$$

where n is the number of electrons (given by the stoichiometry), ε_i is the binding energy of an occupied electronic level i, and d_i is its electronic degeneracy. Note that $\bar{\chi}$ is *not* the same energy as E in Eq. 1. In the helium example, E = IP(He) + IP(He^+) whereas $\bar{\chi}$ = IP(He). Estimating $\bar{\chi}$ of a system (or $\Delta\bar{\chi}$ for a process) requires the combined use of X-ray photoelectron spectroscopy (XPS, for core electrons) and ultraviolet photoelectron spectroscopy (UPS, for valence levels).[40]
The Allen electronegativity scale have demonstrated the feasibility of this approach on single atoms,[43] and it is possible to estimate $\bar{\chi}$ also for larger systems.[40] The methodology is, in principle, straightforward: by knowing the energy of the radiation used, E_{photon}, and the kinetic energy of the ejected electrons, $E_{electron}$, the binding energy, ε_i, of an energy level i can be obtained as,

$$\varepsilon_i = E_{photon} - E_{electron}. \tag{12}$$

However, the interpretation of photoelectron spectra can be challenging when ionization arises from strongly coupled states.[44–46] When analyzing experimental PES spectra, quantum mechanical calculations can therefore be used for comparison and level-identification. Because it is possible to compare with theory, detailed analysis of ionization cross-sections in molecules is not essential; only identification of the main ionization energies is necessary.
Examples: Experimental applications of Eq. 1 can to a large extent rely on literature data of well-characterized molecules and atoms. Table 1 shows an EQC partitioning of four different transformations. Where does these values come from?
The first reaction, which describes H_2 bond formation, requires knowledge of the experimental heat of formation (+2.259 eV) and ionization potential (13.598 eV) of the hydrogen atom, and the bond distance (0.7414 Å), fundamental stretching frequency (4401.2 cm^{-1}) and ionization potential (15.426 eV) of molecular hydrogen. The bond energy, ΔE, can then be calculated according to Eq 10: 0 - 2*2.259 - 0.273 = -4.792 eV, where 0.273 eV is the zero-point energy of H_2. In Table1, all energies are counted in eV per electron, so that $\Delta E/n$ = -4.792/2 = -2.396 eV e^{-1}. $\Delta\bar{\chi}$ for the H_2 bond is calculated as -15.426 - (-13.598) = -1.83 eV e^{-1}, which means that electrons are more strongly bound to H_2 than they are in the H atom. The third term in Eq. 1, the nuclear-repulsion, V_{NN}, is evaluated following Eq. 4. ΔV_{NN} will increase in a transformation where nuclei on average come closer together, and calculates as + 9.716 eV e^{-1} in the formation

of H_2. The last term in Eq. 1, $\Delta E_{ee}/n$ = -10.285 eV e^{-1}, is indirectly obtained following knowledge of the other terms of the equation.

Table 1. EQC energy partitioning of four simple reactions.[a]

Reaction:	$\Delta E/n$	$\Delta\chi,^{-}$	$\Delta V_{NN}/n$	$-\Delta E_{ee}/n$
$2H \rightarrow H_2$, d_{exp} = 0.741 Å	-2.396	-1.828	9.716	-10.285
$H + e^- \rightarrow H^-$ [a]	-0.377	6.045	0.0	-6.422
$H_2 \rightarrow H_2^+ + e^-$, d_{exp} = 1.052 Å	7.870	0.457	-2.872	10.285
$He \rightarrow He^+ + e^-$ [a]	12.293	2.621	0.0	9.672

[a]Experimental data from the NIST Chemistry WebBook. Energies in eV e^{-1}

The EQC-partitioning of the second and third processes listed in Table 1 can similarly be performed knowing the experimental electron affinity (+0.754 eV) of the hydrogen atom, and the molecular bond distance (1.052 Å) and the bond energy (2.651 eV) of H_2^+. Note that some exothermic processes, such as the electron attachment to hydrogen, are *hindered* by a lessened electron binding ($\Delta\chi,^{-}>0$) in the final state. Situations where the $\Delta\chi,^{-}$-term is clearly acting against a favored forward reaction is, in fact, a common occurrence and an indicator attributable to some form of charge flow.[41]

For the last example reaction, helium ionization, it suffices to know the first (24.587 eV) and second (54.418 eV) ionization potential, to calculate ΔE and $\Delta\chi,^{-}$, as was previously mentioned. The most unique property of the EQC-partitioning is that it allows for experimental estimates of electron-electron interaction energies: Because there are no electron-electron interactions in H, H_2^+ or He^+, the ΔE_{ee} -values listed in Table 1 actually provide *absolute* estimate of the electron-electron interactions in H_2, H^- and He. Electron-electron interaction energies are fundamental quantities that underpin all chemical and physical processes and properties, and that previously have only been obtainable from quantum mechanical calculations. In these examples, the obtained electron-electron interaction energies in H_2 (10.285) > He (9.672) > H^- (6.422) eV e^{-1} are, in a way, indirect energy-derived estimates of the relative sizes of these fundamental two-electron-systems.

Further Partitioning: E_{ee} may be further decomposed into the classical electron-electron Coulomb repulsion, E_C, and all remaining non-classical interactions, E_Q, such that $E_{ee} = E_C + E_Q$.[40] This further partitioning is not necessary for the parts of the analysis developed thus far. However, because E_C is calculable provided sufficiently accurate mapping of the electron density, it can offer additional insight into chemical bonding. Experimental determination of electron densities via multipole refinement is a well-established procedure.[22,47,48]

Because the electrostatic terms V_{NN} and E_C both diverge in an infinite crystal they are only strictly attainable for isolated molecular systems, and not for extended materials. However, it is still possible to approximate these terms for molecular solids, if intermolecular interactions are assumed to be small. Omitting the surrounding lattice is common practice when comparing, for

example, crystallographically determined molecular structures with calculations on molecules in a vacuum.[49] In Eq. 1, the terms V_{NN} and $-E_{ee}$ are large and canceling, and for systems with truly extended electronic structures (1D-, 2D-, or 3D-polymers), Eq. 1 is reduced to Eq. 13:

$$\Delta E = n\Delta\bar{\chi} + \Delta(V_{NN} - E_{ee}) \quad \text{(13)}$$

charge relocation energy not directly caused by atomic motion
indirectly calculable

Interpretation and Use of the Average Electron Binding Energy, $\bar{\chi}$

One important term in Eq. 1 is the average electron binding energy, $\bar{\chi}$. Figure 1 demonstrates some of the incantations of $\bar{\chi}$ resolved in energy, real and reciprocal space, in different systems. $\Delta\bar{\chi}$, can be equated to changes in the electronegativity of the system (Fig 1a and 1b).[40, 50] This equality originates from Allen's electronegativity scale,[50] which agrees well (it correlates linearly) with most other scales of electronegativity. Indeed, there are many definitions to electronegativity,[50–55] and the concept has a rich history.[56] Electronegativity is maybe the most important descriptor in chemistry, and it lies at the very heart of chemical rationales and intuition, connecting to several different fields of research.[57, 58] Eq. 1 is an energy partitioning that presents both $\bar{\chi}$ and E_{ee} in clear relation to the total energy E. This means that the relationship shown in Eq. 1 can be used to understand when and why electronegativity arguments work in explaining experimentally observed trends, and what it means when they fail. This ability might prove useful for guiding synthetic chemistry, where such quick rationales are often used.

A related chemically important interpretation of $\bar{\chi}$ is as the average orbital energy, and $n\Delta\bar{\chi}$ as the net orbital stabilization over a transformation. This connects it to ideas of covalency, and molecular orbital theory, so prevalent for the rationalization of chemical bonding. In extended systems, $\bar{\chi}$ also appears in the theoretical framework of moments of the electron distribution, used for describing factors behind solid-state structure (Fig 1c).[59, 60]

$\bar{\chi}$ can be resolved in real space, which is interesting from a QCT perspective (Fig. 1d). When this is done, $\bar{\chi}(\mathbf{r})$ is referred to as the local ionization potential.[61] A one-determinant expression for $\bar{\chi}(\mathbf{r})$ reads as,

$$\bar{\chi}(\boldsymbol{r}) = \frac{\sum_i \rho_i(\boldsymbol{r})\varepsilon_i}{\rho(\boldsymbol{r})}, \quad \text{(14)}$$

where $\rho_i(\mathbf{r})$ is the electronic density of orbital ϕ_i. $\bar{\chi}(\mathbf{r})$ also can be obtained from multideterminant methods.[62] $\bar{\chi}(\mathbf{r})$ plotted on electron density isosurfaces have been used to calculate topological descriptors useful for predicting molecular reactivity,[63, 64] local polarizability, and electronegativity.[58] Experimental determination of molecular orbital densities is becoming possible,[65,66,67] and experimental estimates to $\bar{\chi}(\mathbf{r})$ might be within reach.

The distribution of $\bar{\chi}$ in reciprocal space, $\bar{\chi}(\boldsymbol{k})$, can also be analyzed (Fig. 1e). Whether or not this density can provide insight into the effects of periodicity on chemical bonding in extended

matter is under active study. In the spirit of EQC, $\chi,^{-}(\boldsymbol{k})$ is both theoretically calculable as well as experimentally accessible, using angle-resolved photoelectron spectroscopy (ARPES).

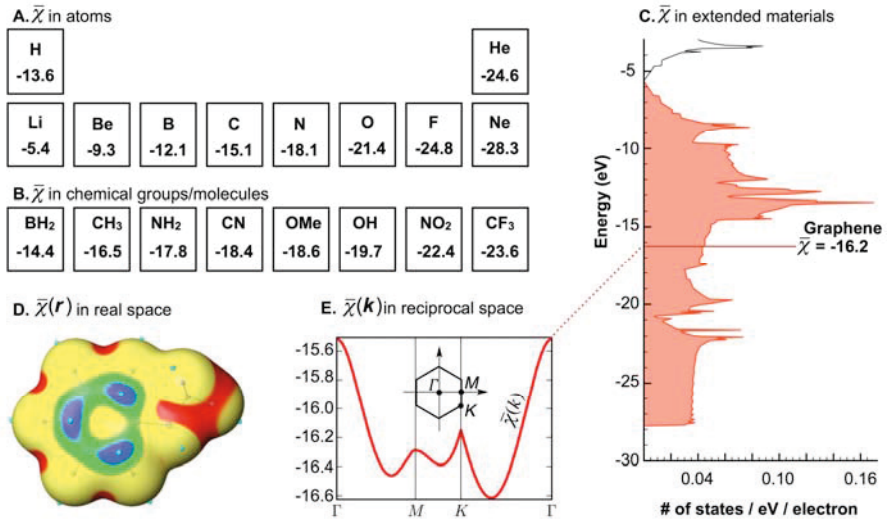


Figure 1. Examples of the average binding energy $\chi,^{-}$ resolved in energy space, real space and reciprocal space. Mention of calculations refer to density functional theory. A) Experimental $\chi,^{-}$ for valence electrons of first and second period atoms obtained from reference 43. B) $\chi,^{-}$ calculated for selected molecules. C) $\chi,^{-}$ calculated for the valence bands of graphene. D) $\chi,^{-}(\mathbf{r})$ calculated on the 0.001 e/bohr^3 isosurface of anisole. Energies (in eV) are denoted as: red > 12.4 > yellow > 9.7 > green > 9.1 > blue. Light blue dots denote local minima on the surface. Reproduced from reference 68 with permission. E) $\chi,^{-}(\boldsymbol{k})$ calculated along the special symmetry points of graphene.

Q – A Descriptor of Chemical and Physical Transformations

Conventional chemical descriptors, such as atomic charge, atomic radii, electronegativity etc., are often essential for molecular design. Such descriptors have in more recent times also become valuable input data for machine learning approaches aimed at high throughput material discovery.[69] Eq. 15 describes Q, one effort to condense the EQC energy partitioning into one value that is descriptive of a given transformation,

$$Q = \frac{n\Delta\bar{\chi} - \Delta(V_{NN} - E_{ee})}{n\Delta\bar{\chi} + \Delta(V_{NN} - E_{ee})} = \frac{n\Delta\bar{\chi} - \Delta(V_{NN} - E_{ee})}{\Delta E} = \frac{2n\Delta\bar{\chi}}{\Delta E} - 1 \qquad (15)$$

Q quantifies the balance between $\Delta\chi,^{-}$, attributable to electronegativity equalization and orbital stabilization, and the combined $\Delta(V_{NN} - \Delta E_{ee})/n$-term, which quantifies charge-drift due to bonding interactions. The Q descriptor is unbound and unit less.

When Q = 1, the $\Delta\chi,^{-}$-term completely describes ΔE, and the situation is attributed to 'perfect covalency'. In contrast, when Q = –1, $\Delta\chi,^{-}$ nets to zero and the $\Delta(V_{NN} - \Delta E_{ee})/n$-term instead completely describes ΔE. The latter situation is associated with 'perfect ionicity'. For energetically favorable interactions where Q>1, electron-electron interactions, described by a

positive $\Delta(V_{NN} - \Delta E_{ee})/n$-term, become more significant. A Q value above 1 signifies increasing degrees of electron correlation effects acting against an energetically favorable event. Similarly, when $Q < -1$, electron-electron interactions play an increasingly important role, but with reversed sign, i.e. the $\Delta(V_{NN} - \Delta E_{ee})/n$-term is negative.

Q has been used to construct a veritable map of bonding interactions and is, when plotted against ΔE for the formation of diatomic molecules, able to distinguish between covalent, electrostatic, dispersion bound, polar, ionic and "metallogenic" (attributed to species that condense to form metals) interactions. This has suggest a tantalizing utility in inferring physical properties of condensed materials by studying smaller systems.[41] *Q* has also proven capable of distinguishing between such subtle effects as red- vs. blue-shifting hydrogen bonds,[70] whose nature is a long standing controversy. Others have implemented Q to study molecular oxides, and concluded it to be a valuable complement in bonding analysis.[37]

Some examples of Q for a representative range of bonding interactions in diatomic molecules are shown in Table 2, alongside the EQC partitioning. Note especially that for bond formations attributable to charge-transfer, polarity, or "metallogenic" character, $\Delta\chi,^{-}$ is destabilizing (positive) and Q is negative.

Table 2. Hybrid theory/experimental EQC-partitioning of diatomic bond formation.[a]

	$\Delta E/n$	$\Delta\chi,^{-}$	$\Delta V_{NN}/n$	$-\Delta E_{ee}/n$	$\Delta(V_{NN} - \Delta E_{ee})/n$	Q
$2\,H \rightarrow H_2$	−2.396	−1.83	+9.72	−10.28	−0.57	0.5
$C + O \rightarrow CO$	−0.806	−2.02	+43.76	−42.54	+1.22	4.0
$H + F \rightarrow HF$	−0.617	+0.42	+14.14	−15.18	−1.04	-2.4
$2\,Li \rightarrow Li_2$	−0.181	+0.54	+8.08	−8.80	−0.72	-6.9
$Na + Cl \rightarrow NaCl$	−0.153	+2.21	+40.74	−43.10	−2.36	-29.9

[a] ΔE and ΔV_{NN} are experimental, $\Delta\chi,^{-}$ is from DFT. Energies are in eV e^{-1}. Q is unit less. Data is reproduced from reference 41 with permission.

Summary and Conclusions

The Experimental Quantum Chemistry (EQC) method, shown as Eq.1, enables the interchangeable use of theoretical calculations and experimental data in a single framework to describe energy processes.[40] The $\Delta\chi,^{-}$ term is conceptually valuable in that it straightforwardly quantifies the average-change of electron binding over a transformation. $\chi,^{-}$ can be resolved in energy, real and reciprocal space, and $\Delta\chi,^{-}$ can be attributed to central chemical concepts such as electronegativity equalization, covalency and orbital stabilization. ΔV_{NN} quantifies changes to the nuclear structure, and ΔE_{ee} summarizes the changes to all forms of electron-electron interactions, including correlation and exchange energies. These EQC terms, including experimental estimates to ΔE_c, and indirect quantification of ΔE_q over a transformation, might offer useful wavefunction constraints and complement other advances in quantum crystallography.

A descriptor Q, defined by Eq. 15, offers an effective simplification for the EQC energy partitioning, and is a straightforward way to analyze the underlying electronic structure changes of chemical or physical transformations on multiple length-scales, calculated or measured. A python script for automated EQC-partitioning and calculation of Q can be found at https://github.com/martinrahm/X-analysis. The script is currently limited to single-reference calculations, and parses the output of several common quantum mechanical software, including Gaussian and ORCA.

References

(1) Fukui, K.; Yonezawa, T.; Shingu, H. *J. Chem. Phys.* **1952**, *20*, 722-725.
(2) Hoffmann, R.; Lipscomb, W. N. *J. Chem. Phys.* **1962**, *36*, 2179-2189.
(3) Woodward, R. B.; Hoffmann, R. *J. Am. Chem. Soc.* **1965**, *87*, 395-397.
(4) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833-1840.
(5) Shaik, S. *New J. Chem.* **2007**, *31*, 2015-2028.
(6) Hoffmann, R.; Shaik, S.; Hiberty, P. C. *Acc. Chem. Res.* **2003**, *36*, 750-756.
(7) Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457-465.
(8) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916-4926.
(9) Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 296-299.
(10) Deringer, V. L.; Tchougreeff, A. L.; Dronskowski, R. *J. Phys. Chem. A* **2011**, *115*, 5461-5466.
(11) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325-340.
(12) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1558-1565.
(13) Francisco, E.; Pendas, A. M.; Blanco, M. A. *J. Chem. Theory Comput.* **2006**, *2*, 90-102.
(14) von, H., Moritz; Frenking, G. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 43-62.
(15) Bickelhaupt, F. M.; Houk, K. N. *Angew. Chem. Int. Ed.* **2017**, *56*, 10070-10086.
(16) Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. *J. Chem. Theory Comput.* **2016**, *12*, 4778-4792.
(17) Thirman, J.; Head-Gordon, M. *J. Phys. Chem. A* **2017**, *121*, 717-728.
(18) Horn, P. R.; Mao, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067-23079.
(19) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *J. Chem. Theory Comput.* **2007**, *3*, 358-374.
(20) Geerlings, P.; De, P., F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793-1873.
(21) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Clarendon Press: Oxford, 1990;
(22) Stokkebro, S., Mette; Overgaard, J.; Brummerstedt, I., Bo. *Z. Anorg. Allg. Chem.* **2013**, *639*, 1922-1932.
(23) Blanco, M. A.; Martin Pendás, A.; Francisco, E. *J. Chem. Theory Comput.* **2005**, *1*, 1096-1109.
(24) Ayers, P. L.; Boyd, R. J.; Bultinck, P.; Caffarel, M.; Carbó-Dorca, R.; Causá, M.; Cioslowski, J.; Contreras-Garcia, J.; Cooper, D. L.; Coppens, P.; Gatti, C.; Grabowsky, S.; Lazzeretti, P.; Macchi, P.; Martín Pendás, Á.; Popelier, P. L. A.; Ruedenberg, K.; Rzepa, H.; Savin, A.; Sax, A.; Schwarz, W. H. E.; Shahbazian, S.; Silvi, B.; Solà, M.; Tsirelson, V. *Computational and Theoretical Chemistry Special Issue: Understanding structure and reactivity from topology and beyond* **2015**, *1053*, 2-16.
(25) Casati, N.; Kleppe, A.; Jephcoat, A. P.; Macchi, P. *Nat. Commun.* **2016**, *7*, 10901pp.

(26) Baranov, A. I.; Ponec, R.; Kohout, M. *Chem. Modell.* **2016**, *12*, 53-83.
(27) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397-5403.
(28) Savin, A.; Jepsen, O.; Flad, J.; Andersen, O. K.; Preuss, H.; Von, S., Hans Georg. *Angew. Chem.* **1992**, *104*, 186-8 (See also Angew. Chem., Int. Ed. Engl., 1992, 31(2), 187.
(29) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683-686.
(30) Kohout, M. *Int. J. Quantum Chem.* **2004**, *97*, 651-658.
(31) Narth, C.; Maroun, Z.; Boto, R. A.; Chaudret, R.; Bonnet, M.-L.; Piquemal, J.-P.; Contreras-Garcia, J. **2016**, 491-527.
(32) Contreras-Garcia, J.; Cardenas, C. *Journal of Molecular Modeling* **2017**, *23*, 271.
(33) Jayatilaka, D.; Grimwood, D. J. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2001**, *A57*, 76-86.
(34) Grabowsky, S.; Genoni, A.; Burgi, H.-B. *Chem. Sci.* **2017**, *8*, 4159-4176.
(35) Dittrich, B. *Acta Crystallographica Section B* **2017**, *73*, 325-329.
(36) Macchi, P. *Acta Crystallographica Section B* **2017**, *73*, 330-336.
(37) Fugel, M.; Beckmann, J.; Jayatilaka, D.; Gibbs, G. V.; Grabowsky, S. *Chemistry* **2018**,
(38) Miao, M.-s.; Hoffmann, R. *Journal of the American Chemical Society*
J. Am. Chem. Soc. **2015**, *137*, 3631-3637.
(39) Cui, Z.-h.; Yang, W.-s.; Zhao, L.; Ding, Y.-h.; Frenking, G. *Angew. Chem., Int. Ed.* **2016**, *55*, 7841-7846.
(40) Rahm, M.; Hoffmann, R. *J. Am. Chem. Soc.* **2015**, *137*, 10282-10291.
(41) Rahm, M.; Hoffmann, R. *J. Am. Chem. Soc.* **2016**, *138*, 3731-3744.
(42) Irikura, K. K. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389-397.
(43) Mann, J. B.; Meek, T. L.; Allen, L. C. *J. Am. Chem. Soc.* **2000**, *122*, 2780-2783.
(44) Deleuze, M. S.; Cederbaum, L. S. *Adv. Quantum Chem.* **1999**, *35*, 77-94.
(45) Cederbaum, L. S.; Domcke, W.; Schirmer, J.; Von Niessen, W. *Adv. Chem. Phys.* **1986**, *65*, 115-159.
(46) Cederbaum, L. S.; Domcke, W.; Schirmer, J.; Von Niessen, W.; Diercksen, G. H. F.; Kraemer, W. P. *J. Chem. Phys.* **1978**, *69*, 1591-1603.
(47) Stokkebro, S., Mette; Bjerg, L.; Overgaard, J.; Krebs, L., Finn; Hellerup, M., Georg Kent; Sugimoto, K.; Takata, M.; Brummerstedt, I., Bo. *Angew. Chem., Int. Ed.* **2013**, *52*, 1503-1506.
(48) Macchi, P.; Gillet, J.-M.; Taulelle, F.; Campo, J.; Claiser, N.; Lecomte, C. *IUCrJ* **2015**, *2*, 441-451.
(49) Becker, S.; Müller, P. *Chem. Eur. J.* **2017**, *23*, 7081-7086.
(50) Allen, L. C. *J. Am. Chem. Soc.* **1989**, *111*, 9003-9014.
(51) Pauling, L. *J. Am. Chem. Soc.* **1932**, *54*, 3570-3582.
(52) Mulliken, R. S. *J. Chem. Phys.* **1934**, *2*, 782-793.
(53) Allred, A. L.; Rochow, E. G. *J. Inorg. Nucl. Chem.* **1958**, *5*, 264-268.
(54) Hinze, J.; Jaffe, H. H. *J. Am. Chem. Soc.* **1962**, *84*, 540-546.
(55) Pearson, R. G. *J. Am. Chem. Soc.* **1985**, *107*, 6801-6806.
(56) Jensen, W. B. *J. Chem. Educ.* **1996**, *73*, 11-20.
(57) Grochala, W.; Hoffmann, R. *New J. Chem.* **2001**, *25*, 108-115.
(58) Politzer, P.; Peralta-Inga, S., Zenaida; Bulat, F. A.; Murray, J. S. *J. Chem. Theory Comput.* **2011**, *7*, 377-384.

(59) Burdett, J. K.; Lee, S. *J. Am. Chem. Soc.* **1985**, *107*, 3063-3082.
(60) Pettifor, D. *Bonding and Structure of Molecules and Solids*; Oxford University Press: New York, 1995;
(61) Jin, P.; Murray, J. S.; Politzer, P. *Int. J. Quantum Chem.* **2004**, *96*, 394-401.
(62) Ryabinkin, I. G.; Staroverov, V. N. *J. Chem. Phys.* **2014**, *141*, 084107/1-084107/8.
(63) Murray, J. S.; Shields, Z. P.-I.; Lane, P.; Macaveiu, L.; Bulat, F. A. *J. Mol. Model.* **2013**, *19*, 2825-2833.
(64) Toro-Labbe, A.; Jaque, P.; Murray, J. S.; Politzer, P. *Chem. Phys. Lett.* **2005**, *407*, 143-146.
(65) Itatani, J.; Levesque, J.; Zeidler, D.; Niikura, H.; Pepin, H.; Kieffer, J. C.; Corkum, P. B.; Villeneuve, D. M. *Nature (London, U. K.)* **2004**, *432*, 867-871.
(66) Wiessner, M.; Hauschild, D.; Sauer, C.; Feyer, V.; Schoell, A.; Reinert, F. *Nat. Commun.* **2014**, *5*, 4156.
(67) Puschnig, P.; Boese, A. D.; Willenbockel, M.; Meyer, M.; Lueftner, D.; Reinisch, E. M.; Ules, T.; Koller, G.; Soubatch, S.; Ramsey, M. G.; Tautz, F. S. *J. Phys. Chem. Lett.* **2017**, *8*, 208-213.
(68) Politzer, P.; Murray, J. S.; Bulat, F. A. *J. Mol. Model.* **2010**, *16*, 1731-1742.
(69) Raccuglia, P.; Elbert, K. C.; Adler, P. D. F.; Falk, C.; Wenny, M. B.; Mollo, A.; Zeller, M.; Friedler, S. A.; Schrier, J.; Norquist, A. J. *Nature* **2016**, *533*, 73-76.
(70) Sessler, C. D.; Rahm, M.; Becker, S.; Goldberg, J. M.; Wang, F.; Lippard, S. J. *J. Am. Chem. Soc.* **2017**, *139*, 9325-9332.

Crystal Structure Refinement

Peter Müller
Massachusetts Institute of Technology, Chemistry Dept, Cambridge, MA, USA
pmueller@mit.edu

Introduction

Refinement is the process of iterative alteration of the molecular model with the goal to maximize its compliance with the diffraction data" (Müller, 2009). The structure solution from direct methods, charge flipping, intrinsic phasing, or Patterson methods is frequently already fairly good. However, the atomic coordinates are not quite accurate, the atom types of some atoms have been assigned incorrectly (if at all), and details of the structure are missing (hydrogen atoms, disorders, solvent molecules, *etc.*).

It is important to note that the atomic positions in the first solution, sometimes called the initial model or the *trial structure*, are not the direct result of the diffraction experiment but an interpretation of the electron density function calculated from the measured intensities and the somehow determined phase angles.

Critically assessing the initial model and making sensible changes usually affords an improved structural model and better phases can be calculated from the atomic positions of this improved trial structure. This, in turn, allows re-determination of the electron density function with higher accuracy. From the new, more accurate electron density map, an even better model can be derived, which leads to even better phase angles, and so forth. In every such cycle, adjustments to the atomic model are made: atom types are changed, missing atoms are introduced, *etc.* This iterative process is called structure refinement.

Generally speaking, structure refinement consists of three major steps: First, close examination of the

F_o-F_c **map** (see below) helps to introduce new atoms and remove "bad" ones. Second, when all non-hydrogen atoms are found, the atoms can be refined anisotropically. And third, once the model is anisotropic, the hydrogen atom positions can be determined or calculated.

Electron Density Maps

Simply put, the Fourier transform of the diffraction pattern is the three-dimensional electron density function, which can be expressed in terms of electrons per cubic Ångstrom and graphically represented in different ways. Crystallographers rely on several so called electron density maps, the most important ones are the following three:

***F_o* map**: Electron density calculated using observed structure factures combined with phases calculated from the atomic model. This map shows the observed electron density; its accuracy depends largely on the accuracy of the phases.

***F_c* map**: Electron density calculated from structure factors and phases calculated from the atomic model. This map shows the electron density according to the model only.

***F_o*-*F_c* map**: The difference between the two other maps (model subtracted from experimental density). This map exhibits values close-to-zero for parts of the structure where the model is consistent with the experimental density, it shows large positive values at places where the model should have an atom, but does not, and large negative values at places where the model has an atom that should not be there. Weaker positive or negative values for the F_o-F_c map could point to wrongly assigned atom types. The F_o-F_c map is often called difference density map and is the one with the most practical relevance of the three.

Figure 1 shows part of a trial structure, namely a Cp ring, alongside the corresponding region of the F_o and F_o-F_c maps. Examination of the electron density maps shows that the ligand in the current model is not actually a Cp ring but, in fact, a Cp* ring.

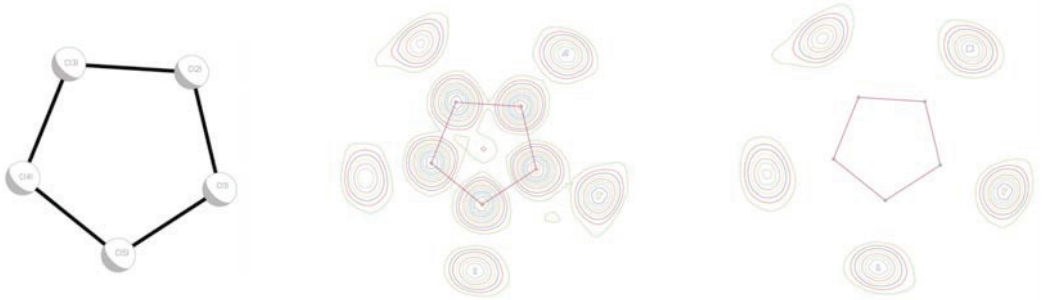


Figure 1. Left: Trial structure (a Cp ring). Middle: Trial structure with superimposed F_o map. Right: Trial structure with superimposed F_o-F_c map. Interpretation of the residual electron density as hydrogen atoms is not sensible as (a) the density is too high and (b) the location of the electron density maxima is too far from the carbon atoms of the Cp ring.

Anisotropic Displacement Parameters

Once all non-hydrogen atoms have been assigned correctly, the structure can be refined anisotropically. It is reasonable to assume that atoms move with different amplitudes in different directions. Instead of describing an atom as a sphere, it therefore is described as an ellipsoid; the elements of the 3X3 matrix defining the ellipsoid are called anisotropic displacement parameters (ADPs). Since the ADP-matrix, usually called the U^{ij}-matrix, is symmetrical, this adds only six parameters for each anisotropic atom, not nine. Size and shape of the thermal ellipsoids are excellent indicators for problems with the crystallographic model. Figure 2 shows a Cp* ring with both isotropic and anisotropic modeling.

The representation of the Cp* ring shown in Figure 2 is called a thermal ellipsoid plot (sometimes also ORTEP-style plot, named after Carroll Johnson's Oak Ridge Thermal Ellipsoid

Plot software) and individual atoms are represented by ellipsoids that indicate the magnitude and direction of the thermal vibration of each non-hydrogen atom in the molecule. The thermal ellipsoids are scaled to include the space occupied by a subset (typically 50%) of the electrons associated with the atoms they represent. Significantly too large or too small thermal ellipsoid volumes indicate incorrect atom type assignment. For easier comparison, the volume of a thermal ellipsoid can be represented by a single numerical value, the *U*-equivalent value (U_{eq}). The U_{eq} is defined as one third of the trace of the orthogonalized matrix U^{ij} describing the anisotropic displacement-ellipsoid.

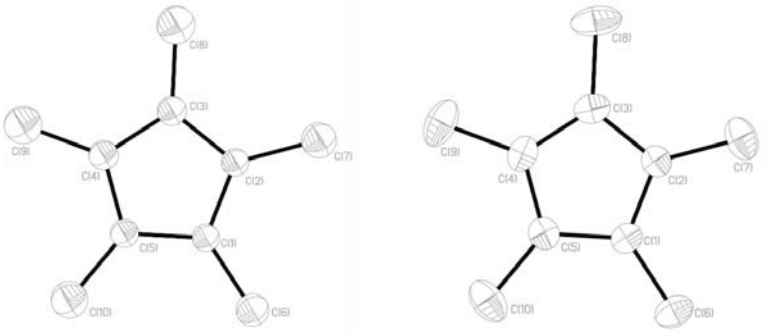


Figure 2. Isotropic (left) and anisotropic (right) refinement of the carbon atoms in a Cp* ring.

In theory, all atoms within a molecule should have roughly the same thermal ellipsoid size, and thermal ellipsoids should not be strongly prolate or oblate; however some atoms in a molecule can move more freely than other atoms (like terminal groups or moieties that can rotate about one of the bonds). Therefore, terminal atoms generally show somewhat larger and possibly somewhat more elongated ellipsoids when compared to atoms that are, for example, part of a cyclic moiety or the central atom in an organometallic complex. Strongly elongated thermal ellipsoids may indicate disorder (see Figure 3).

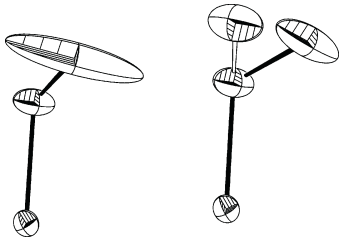


Figure 3. Anisotropic displacement parameters of a disordered ethyl group; on the left without and on the right with modeling of the disorder (empty lines for the minor component). When disorder is ignored, the refinement software tries to describe both atom positions with one elongated ellipsoid. Figure reproduced from Müller, 2006.

If the thermal ellipsoid (or the corresponding U_{eq} value) for a given atom in a molecular model is much smaller or larger than those of similarly positioned atoms, the atom in question may have been incorrectly assigned (*e.g.*, nitrogen versus carbon). As mentioned above, thermal ellipsoids

are typically scaled to represent 50% electron density. That means that, for example, the thermal ellipsoid of an oxygen atom in a molecular model contains four electrons, the ellipsoid of a carbon three. If an oxygen atom is incorrectly assigned the element type carbon, the resulting thermal ellipsoid will be too small, as the volume containing three electrons in an oxygen atom is smaller than that containing four electrons. Figure 4 shows a tetrahydrofuran molecule where all five atoms were modelled as carbon; it is easy to identify which of the five atoms is the oxygen.

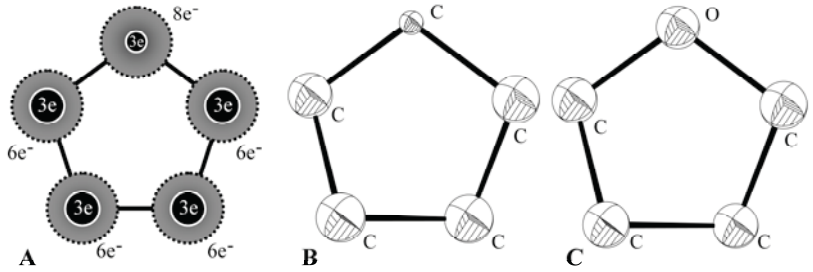


Figure 4. A: Cartoon of a tetrahydrofuran (THF) molecule. The dashed circles in grey represent the atoms, the white rimmed black circles the volume increment corresponding to three electrons. **B**: Isotropic displacement parameters at the 50% level of a THF molecule where all five atoms were refined as carbon; the sphere representing the oxygen atom is much smaller than the other spheres. **C**: The same THF molecule with correct atom type assignment; all spheres exhibit approximately the same volume. Figure reproduced from Müller, 2006.

A real-live case where thermal ellipsoids were used to identify the nitrogen and oxygen atoms in the structure of a natural compound (isolated and purified from nature, not synthesized in a lab) is shown in Figure 5. Refining all atoms as carbon makes the heavier, more electron rich nitrogen and oxygen atoms exhibit smaller thermal ellipsoids.

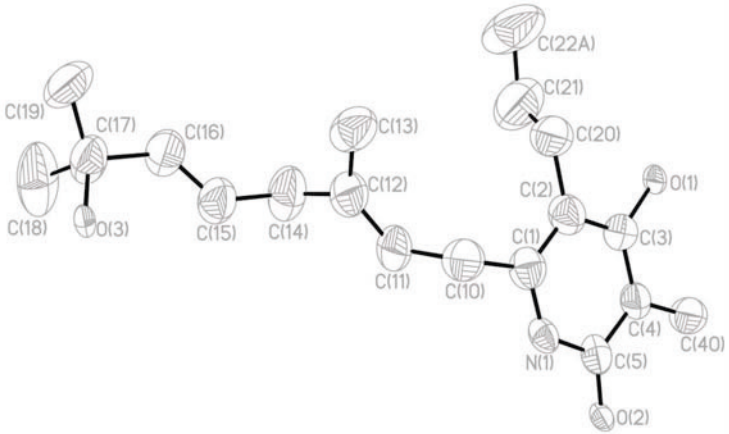


Figure 5. Thermal ellipsoids can help to identify atom types: Even though the atoms have the correct atom names in the figure above, all atoms in the model were refined as carbon atoms. As a result, the thermal ellipsoids for atoms that are, in fact, nitrogen and oxygen are significantly smaller than those of actual carbon atoms.

Thermal motion of atoms in the crystal can also lead to large ADPs. The very large ellipsoids in the anion $N(SO_2F)_2^-$ at room temperature can be explained either with strong atomic movement or disorder (Figure 6). In this case, the generally preferred staggered conformation at 112K and the ecliptic conformation at 290K make this look more like a disorder. This is corroborated by the bond lengths.

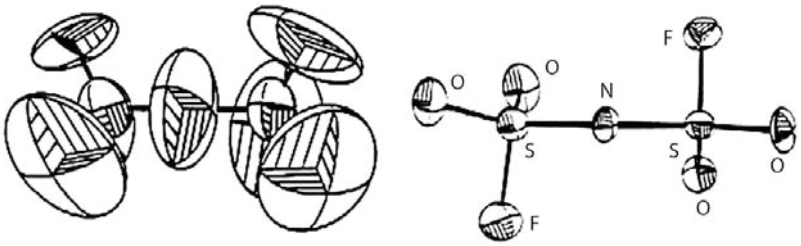


Figure 6. Thermal ellipsoid representation of the same structure at 290K (left) and 112K (right). At 290K all S-O and S-F distances are similar and between 1.24 and 1.36 Å. At 112K, the different bonds can be easily distinguished: S-F: 1.57, and S=O: 1.42 Å. Figure courtesy of George M. Sheldrick.

Least-Squares Approach

The diffraction experiment gives us intensities, which correspond to squared structure factor amplitudes ($I \sim F^2$). Structure factors F are complex numbers, *i.e.* vectors in the Argand plane:

$$F(hkl) = \sum_i f_i \left[\cos 2\pi\left(hx_i + ky_i + lz_i\right) + i \sin 2\pi\left(hx_i + ky_i + lz_i\right)\right]$$

Summation over all atoms *i*. *F*: structure factor; *f*: atomic scattering factor; *h,k,l*: Miller indices; *x,y,z*: atom coordinates.

By means of Fourier transformation, a complete set of structure factors is calculated from the atomic model. The calculated intensities are then compared with the measured intensities, and the best model is that, which gives the smallest value for the minimization function M.

$$M = \sum w\left(F_o^2 - F_c^2\right)^2$$

F: structure factor; *o*: observed; *c*: calculated; *w* weighting factor (derived from σ).

In order to find the minimum, the first derivatives of M have to be set equal to zero with respect to each parameter, resulting in one equation per parameter to be refined. Unfortunately, those equations are not linear (they contain trigonometry terms and exponential functions); however given a reasonably good starting model (*i.e.* structure solution) one can calculate **shifts** of parameter values rather than the parameters themselves. This is done from a set of linear equations (Taylor series about all starting model parameters using only the first-derivative terms).

$$\Delta F_c = \frac{\partial F_c}{\partial x_1}\Delta x_1 + \frac{\partial F_c}{\partial y_1}\Delta y_1 + \ldots$$

This is acceptable under the assumption that the required shifts are so small that the higher terms are negligible. Convergence is slow (usually needs 6 to 10 cycles or more) but with every cycle the shifts become smaller and the approximation holds better.
If the starting model is not good enough, the required shifts may be too large and the method may not converge to the correct minimum. Damping limits the amount of shift for example using the Marquardt algorithm or by scaling all shifts by a number <1 (for example 0.7). Constraints and restraints also stabilize the refinement (see below).

Quality of the Model

The quality of a crystallographic model can be assessed by several figures of merit. The best criterion is probably the standard uncertainties of the bond distances; most commonly, however, the residual values *wR* and *R*. Another commonly used figure of merit is the Goodness of Fit (GoF or GooF or simply *S*).

$$wR = \left[\frac{\sum w\left(F_o^2 - F_c^2\right)^2}{\sum wF_o^2} \right]^{1/2}$$

*wR*2: Most closely related to refinement against F^2.

$$R = \frac{\sum \left| |F_o| - |F_c| \right|}{\sum |F_o|}$$

R or *R*1: Most popular *R*-value, based on *F*.

$$S = \left[\frac{\sum w\left(F_o^2 - F_c^2\right)^2}{\left(N_R - N_P\right)} \right]^{1/2}$$

GooF: *S* is supposed to be > 1.0

F: structure factor; *o*: observed; *c*: calculated; *w* weighting factor (derived from σ). N_R: number of independent reflections; N_P: number of refined parameters.

Parameters

A parameter is any quantity that is adjusted during structure refinement. For every generally located atom in the model, there are three positional coordinates (*x*, *y*, *z*) and six anisotropic displacement parameters to be refined (only one displacement parameter for isotropic atoms). In addition, there is one overall scale factor per structure and possibly several additional factors, like a twin ratio for twinned structures, a Flack-parameter for non-centrosymmetric structures (Flack, 1983), a parameter for extinction, *etc.* Some refinement programs like SHELXL (Sheldrick, 2015) allow for additional free variables to be refined that can be assigned to various parameters like occupancy factors or interatomic distances. Altogether, the number of parameters is roughly nine to ten times the number of independent atoms in a structure. The International Union of Crystallography (IUCr) recommends for a stable refinement a minimum data-to-parameter-ratio of > 8 for non-centrosymmetric structures and > 10 for centrosymmetric structures. This corresponds to a resolution of *ca.* 0.84 Å or a 2Θ angle of 50°

(Mo) or 135° (Cu). Constraints and restraints improve the data-to-parameter-ratio: constraints remove parameters, restraints add data.

Constraints

Constraints are mathematical equations, rigidly relating two or more parameters or assigning fixed numerical values to certain parameters, hence reducing the number of independent parameters to be refined. Site occupancy factors are constraints present in almost every structure. Even for disordered atoms the sum of the occupancies is usually constrained to add up to 1.0. Another typical constraint is the rigid group. Atoms within a rigid group are refined as a unit that does not change its shape and only translation and rotation of the group as a whole are refined. Hydrogen atoms can be "placed on mathematically calculated positions and refined using a riding model". That means X-H *distances* and H-X-H or H-X-Y *angles* are constraint to certain values, not the actual hydrogen *positions*!

Atoms on special positions require constraints for their coordinates, occupancies and sometimes also their ADPs. Figure 7 summarizes the special position constraints to be applied to an atom residing on the *b*-axis (for example in the monoclinic space group *P*2)

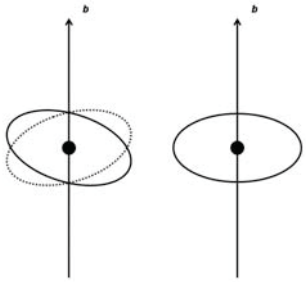


Figure 7. Atom located on a twofold axis along *b*. A 180° rotation about *b* must not change the position of the atom or the shape of the thermal ellipsoid. From the first condition follows (x, y, z) = (-x, y, -z), which is true only for x = z = 0. The second condition dictates: (U^{11}, U^{22}, U^{33}, U^{23}, U^{13}, U^{12}) = (U^{11}, U^{22}, U^{33}, -U^{23}, U^{13}, -U^{12}), which is only true for U^{23} = U^{12} = 0. The left-hand side of the figure shows an incorrectly shaped thermal ellipsoid mapped onto itself by the twofold; the right-hand side shows a correctly shaped one. Figure reproduced from Müller, 2006.

Restraints

Restraints are assumptions used to introduce chemical or physical information into a refinement as additional experimental observations. Restraints are treated as data (with a standard uncertainty, also called elasticity) and in the presence of restraints the minimization function introduced above changes as follows:

$$M = \sum w\left(F_o^2 - F_c^2\right)^2 + \sum 1/\sigma^2\left(R_t - R_o\right)^2$$

Minimization Function including restraints: *F*: structure factor; *o*: observed; *c*: calculated; *w* weighting factor; σ: standard uncertainty assigned to the restraint; R_t: target value for restraint quantity; R_o: actual value as observed in the molecular model.

It is apparent that restraints are treated exactly like data against which the parameters are refined. In many refinements, restraints may not be needed at all. However, when the data-to-

parameter ratio is low, or when correlations among certain parameters occur (*e.g.* for the refinement of disorders, pseudo-symmetry or twinned structures), restraints can become essential. "Restraints should be used with great care and only if justified. When appropriate, however, they should be used without hesitation, and having more restraints than parameters in a refinement is nothing to be ashamed of." (Müller, 2009)

In general, one can distinguish two different types of restraints: direct restraints and relative restraints. Direct restraints assign outside values to certain parameters of a model while relative restraints, which are also called similarity restraints, relate equivalent parameters within a model. The most important restraints are geometrical restraints and restraints on anisotropic displacement parameters (ADPs); both geometry- and ADP-restraints can be either direct or relative in nature. To include, for example, information about carbon–carbon single bonds into a refinement, one could either assign a sensible target value to all C—C bonds (say 1.54 Å) or one could specify that all C—C bonds in the model should have approximately the same length whatever that value may be. The first approach describes a direct restraint taking into account results from spectroscopy and/or other sources like databases, the second approach is that of similarity restraints. The advantage of relative restraints is that there is no need for "outside" information and also that the refinement converges well. The disadvantage of relative restrains lies in potentially underestimated standard uncertainties of bond lengths and angles, especially when many restraints are used. In general, relative restraints can be considered milder and, whenever possible, should be given a preference over direct restraints. Figure 8 shows geometry restraints applied to Cp*.

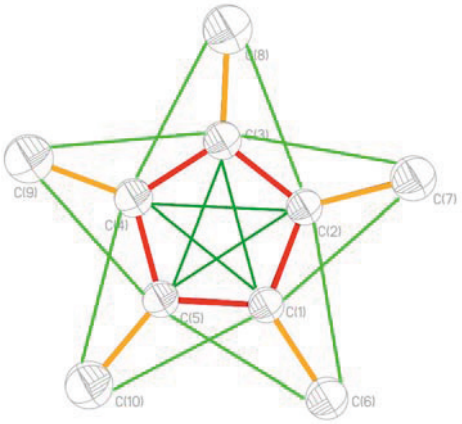


Figure 8. Geometry restraints for a Cp* ring. All five bonds of the Cp-core are approximately the same (red lines), all C—C single bonds are approximately the same (orange lines), the five C-C-C-angles in the Cp-core are all similar (dark green lines) and the ten C-C-C-angles formed by the methyl groups relative to the Cp-core are similar (light green lines). SHELXL treats angular restraints as restraints on 1,3-distances. Typical standard uncertainties are 0.02 Å for 1,2-distance restraints and 0.04 Å for 1,3-distance restraints In addition to the distance restraints one may assume that all ten atoms of the Cp* ligand reside on one common plane (typically within 0.1 $Å^3$).

It is fair to assume that atoms bound to one another move similarly, both in direction and amount. One can restrain the anisotropic displacement parameters of two atoms *in the direction of the bond* between them to be equal within a given standard uncertainty (*e.g.* 0.01). This is called a "rigid bond restraint". It can also be assumed that atoms that are near one another in a structure move in similar directions with approximately similar amplitudes. Therefore, one can restrain atoms close to one another to have the same U^{ij} components within a given standard uncertainty (*e.g.* 0.04). This is called a "similar-ADP restraint". The second assumption is much bolder then the first one, hence the much larger standard uncertainty. In addition, one can restrain anisotropically refined atoms to behave approximately isotropically within a given standard uncertainty (say 0.1 $Å^2$). Figure 9 summarizes commonly used ADP-restraints.

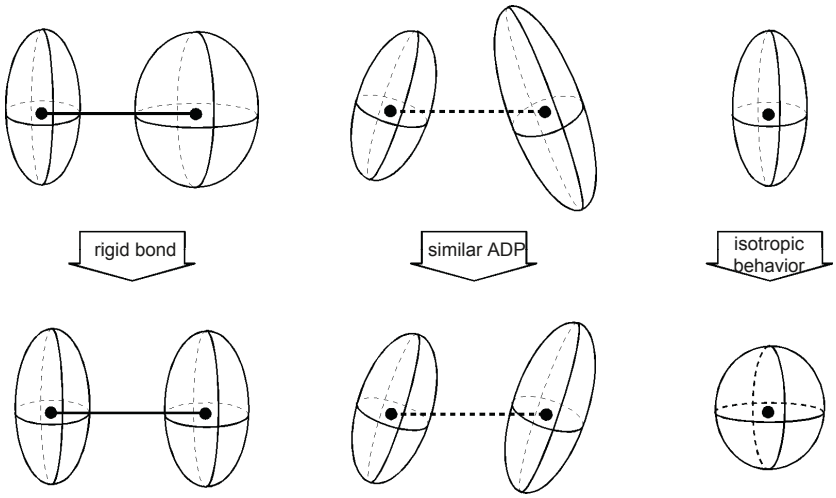
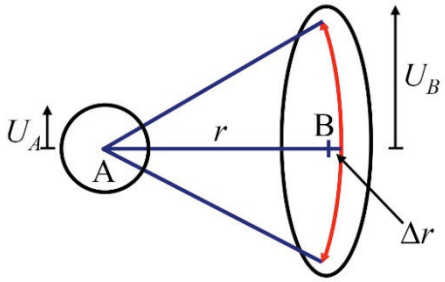


Figure 9. Common ADP restraints. Figure courtesy of Thomas R. Schneider.

Libration

In X-ray and electron crystallography, there is a conterintuitive effect that lets bond distances *appear* shorter at higher temperatures, even though the unit cell gets slightly larger at the same time. The reason for this is libration. Especially terminal atoms show approximately circular motion with the bond as radius. In anisotropic refinement, this motion is fitted as an ellipsoid, the center of which lies inside the circle of motion, while the atom is located on its periphery. This effect makes the bond distance appear shorter than it actually is (see Figure 10). Depending on the movement of the atoms, the correction that needs to be applied is somewhere between 0.001 – 0.1 Å. Libration is much stronger at higher temperatures and particularly pronounced for hydrogen atoms, small ions, like NO_3^-, BF_4^-, ClO_4^-, PF_6^-, and for $-CF_3$ groups. For most low-temperature structures, the effects of libration are smaller than the standard uncertainties of the bond lengths and it is not necessary to correct for it, however a simple equation estimating the librational bond shortening as a function of atomic motion is given as part of Figure 10.



$$\Delta r \approx \frac{\Delta U}{2r} = \frac{[U_B - U_A]}{2r}$$

Figure 10. Apparent bond shortening due to libration in the example of two atoms, A and B. The average displacements of atoms A and B are described by U_A and U_B and the ellipsoids fitting the atomic motion are drawn as a black line (almost circular for atom A and significantly elongated for atom B). The calculated distance r is too short by the length Δr. Δr can be estimated according to the equation on the left-hand side. Figure courtesy of George M. Sheldrick.

Hydrogen Atoms

Once all non-hydrogen atoms are refined anisotropically, hydrogen atoms can be introduced. Hydrogen atoms interact with the x-ray or electron beam only comparatively weakly and, especially in the presence of heavy atoms, hydrogen atoms are notoriously difficult to detect with X-ray and electron diffraction methods. As a result, their location can be determined less accurately. Most notably, the relatively high electron density between the atoms and libration effects make X—H bonds appear too short (Figure 11).

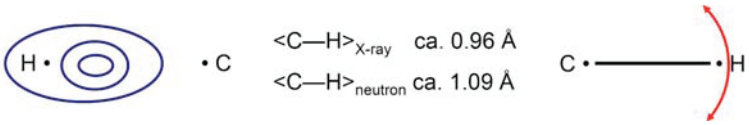


Figure 11. Relatively high electron density between hydrogen and carbon (or whichever other atom the hydrogen atom may bind to) as well as libration make bond distances involving hydrogen appear significantly shorter based on X-ray and electron diffraction data when compared to neutron diffraction. Figure courtesy of George M. Sheldrick.

Considering the difficulty of determining hydrogen positions with accuracy, it is customary to include hydrogen atoms into their geometrically calculated positions. In most cases, the positioning of hydrogen atoms bound to carbon in an atomic model during the refinement of an X-ray crystal structure is done entirely without any or only very little direct information from the diffraction experiment, as the standard bond lengths and angles are well known. Hydrogen atoms on aromatic carbons, CH, CH_2 groups, and most CH_3 groups are straightforward. Only the torsion angle of CH_3 in acetonitrile, Cp*, toluene, *etc.* needs to be determined, as staggering the hydrogen atom geometry is not an option in such cases (Figure 12).

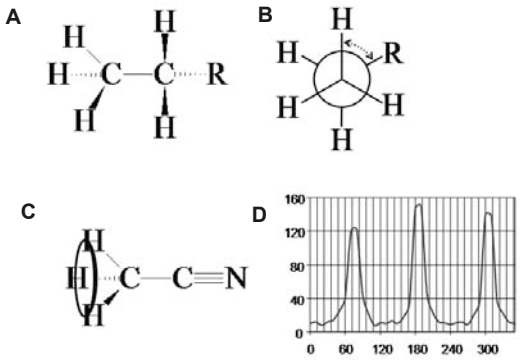


Figure 12. Example of hydrogen atom placement on methyl groups. **A:** The hydrogen positions in an ethyl group can be calculated from the carbon atoms alone (assuming a staggered arrangement). **B:** The same situation as in **A**, but in Newman projection, showing the torsion angle as a dotted line. **C:** Methyl group in acetonitrile: the circle through the hydrogen atoms corresponds to the line in space on which the hydrogen atoms must lie. **D:** Electron density along this circle (simulated data on arbitrary scale the horizontal axis gives the place on the circle in degrees from an arbitrary starting point): expected are three maxima, 120° apart, which correspond to the location of the hydrogen atoms on the circle in **C**. Figure reproduced from Müller, 2006.

After their introduction, the hydrogen atoms are then refined using a riding model (see section about constrains above). The riding model has the advantage that no additional parameters need to be refined when hydrogen atoms are introduced in this fashion (except perhaps torsion angles for methyl groups). On the other hand, potentially acidic hydrogen atoms (those bound to nitrogen or especially oxygen) and hydrogen atoms that are important for the chemistry of the molecule at hand can and should still be included into the model from electron density maxima and then refined semi-freely with the help of distance restraints. Hydrogen atoms of water molecules must be detected in the experimental electron density or else they cannot be included into the model, although sometimes one can infer from the surrounding potential hydrogen bonding partners where the water hydrogen atoms might be.

Even more difficult to detect can be hydrogen atoms in heavy metal hydrides. The sometimes relatively strong Fourier truncation ripples close to heavy atom positions can overpower the rather weak electron density maxima representing the hydrogen atoms. Exceptionally accurate and especially complete high quality data and proper scaling are required to distinguish those hydrogen atoms from the background noise.

Final Remark

The pages above describe the course of a typical crystal structure refinement based on X-ray diffraction data. For electron diffraction, the same principles apply, however the data quality is usually lower. Most notably, the error distribution is less Gaussian for electron diffraction data and structure refinements don't converge as readily as they do for structures based on X-ray

data. That means electron-crystallographers may need to rely more heavily on constraints and restraints, and dampening the refinement is frequently helpful to facilitate convergence.

References

Flack, H. D. (1983). *Acta Cryst.*, **A39**, 876–881.

Müller, P., ed. (2006). *Crystal Structure Refinement*, IUCr Texts on Crystallography Volume 8, Oxford University Press.

Müller, P. (2009). *Crystallography Reviews*, **15**, 57-83.

Sheldrick, G. M. (2015). *Acta Cryst.*, **C71**, 3–8.

QCBED – A Nexus Between Quantum and Electron Crystallography

Philip Nakashima
Dept of Materials Science and Engineering, Monash Univ., Clayton, Victoria, Australia
philip.nakashima@monash.edu

Introduction

All materials properties (with the exception of radioactivity) are largely governed by the distribution of electrons in and between the constituent atoms. Quantum crystallography is the science of extracting quantum mechanically valid information about these electron distributions from diffraction experiments [1]. In other words, electron distributions and chemical bonding and how these are measured and modelled are the central agendas for quantum crystallographers.

The original definition of quantum crystallography referred only to X-ray diffraction experiments. Whilst X-rays are diffracted by their interaction with the electron density around atoms in crystals (see figure 1), electrons are scattered by the crystal potential, which is directly related to the electron density by the Mott formula [2] (figure 1). Therefore, electron diffraction experiments are now considered part of the field of quantum crystallography [3, 4] because they are another means by which to extract quantum mechanically valid information about atoms and bonding in crystals.

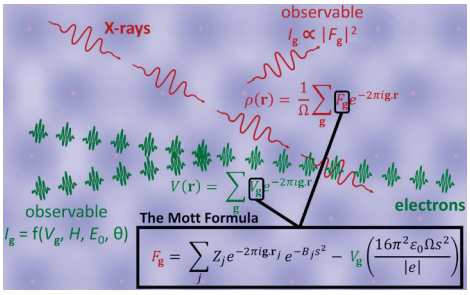


Figure 1. Schematic illustration of the interaction of X-rays with the periodic electron density in a crystal and electrons with the periodic crystal potential. The relationship between the Fourier coefficients (structure factors) of the electron density and those of the crystal potential is expressed by the Mott formula [2], which is derived from Poisson's equation. The electron density structure factors are expressed as $F_\mathbf{g}$, while $V_\mathbf{g}$ denote the crystal potential structure factors. The reciprocal lattice vectors are given by $\mathbf{g}$, whilst $\mathbf{r}$ is a real space position in the unit cell. The subscript j in the Mott formula signifies the j^{th} atom in the unit cell, B_j is the corresponding Debye-Waller factor, Ω is the volume of the unit cell, ε_0 is the free space permittivity, $|e|$ is the magnitude of the charge of an electron and $s = (\sin\theta)/\lambda$ (where θ is the scattering angle and λ is the wavelength of the radiation being used). A single (kinematic) scattering approximation for interpreting X-ray diffraction intensities, $I_\mathbf{g}$, makes these observables much simpler to translate into measured structure factors that in the case of electron diffraction intensities as the latter require a full dynamical scattering treatment because the intensities are complicated functions of not only the structure factors, $V_\mathbf{g}$, but also the specimen thickness (H), the electron energy (E_0) and the scattering angle, θ, i.e. $I_\mathbf{g} = f(V_\mathbf{g}, H, E_0, \theta)$.

A holistic view of quantum crystallography does not just involve the ability of diffraction experiments to furnish us with quantum mechanically valid descriptions of atoms and electrons in crystalline materials. Our ability to interpret and extract meaningful information from diffraction patterns depends on quantum mechanically valid descriptions of the interaction of radiation (X-rays, γ-rays or electrons) with matter. In the case of X-ray or γ-ray diffraction experiments, the application of the kinematic (single) scattering approximation greatly simplifies the analysis of diffracted intensities. However, the validity of such an approximation depends on the nature of the interaction of these radiations with matter and this is in turn based on quantum mechanically valid descriptions of these interactions. Electrons, being charged, interact with matter between 4 and 5 orders of magnitude more strongly than X-rays and γ-rays. This means that a kinematic approximation is almost never appropriate when interpreting electron diffraction patterns. An accurate treatment of the diffracted intensities in electron diffraction patterns therefore requires a full description of dynamical scattering that is based on quantum mechanically valid formulations of the interaction of electrons with the crystal potential. Continuing with a holistic view of quantum crystallography, the precision and accuracy of diffraction experiments is now at a sufficiently high level to allow experimentally measured Fourier coefficients of the crystal potential or electron density (structure factors) to test the validity and ranges of applicability of solid-state theories, such as density functional theory (DFT) for example. These solid-state theories are themselves the constructs of quantum mechanical models and approximations of variable validity.

Table 1. A comparison of the characteristics of X-rays / γ-rays with those of electrons relative to quantum crystallographically relevant diffraction experiments.

X-rays / γ-rays	**Electrons**
• Massless	• Have mass
• Chargeless	• Have charge
• Not easily manipulated with optical elements and cannot be energised but can be monochromated	• Easily manipulated with electromagnetic optical elements, can be energised (accelerated and decelerated) and monochromated before AND after interaction with the specimen
• Typical specimen volumes irradiated ~10^{-15} m^3	• Typical specimen volumes irradiated ~10^{-25} m^3
• Spatial selectivity generally >10^{-8} m	• Spatial selectivity 10^{-11} – 10^{-9} m
• Interact with the electron density	• Interact with the crystal potential
• Kinematic scattering approximation often valid in interpreting data	• Full dynamical scattering description must be applied to interpret data
• Extinction (due to multiple scattering) and scale (due to the relative nature of the integrated intensities) can detract from the accuracy of structure factor measurements [5 – 8].	• The full dynamical (multiple) scattering analysis makes extinction irrelevant. The rocking curve intensities within CBED patterns makes structure factor measurements absolute.

Table 1 gives a comparison of the characteristics of X-rays (and γ-rays) with electrons in the context of diffraction experiments that are relevant to quantum crystallography. The fundamental differences between electrons and photons (X-rays and γ-rays) are that electrons

have mass and charge. This makes them easily manipulable with electromagnetic optical elements and means that they can be accelerated and decelerated by electric fields. As a result, their wavelengths can be easily tuned and they can be focussed into probes with sub-nanometre dimensions even in standard transmission electron microscopes (TEMs). In TEMs with aberration correctors, probe sizes are routinely sub-Ångström. Electromagnetic deflection of incident electron beams with sub-Ångström spatial selectivity, coupled with the ability to image materials and the location for a diffraction experiment at atomic resolution means that regions of perfect crystal can be selected for diffraction experiments, even in materials with nano-crystalline phases and high crystalline defect densities.

In addition to this high level of two-dimensional spatial selectivity, the strong interaction of electrons with matter due to their charge means that typical specimen thicknesses for electron diffraction experiments are nearly always less than about 200 nm. Standard probe sizes used in quantitative convergent-beam electron diffraction (QCBED) with the aim of measuring structure factors of crystal potential are about 1 nm. The volume of irradiated material is therefore of the order of 10^{-25} m^3, i.e. about 10 orders of magnitude smaller than in the case of typical single-crystal X-ray or γ-ray diffraction experiments. This is an effective measure of the differences in three-dimensional spatial selectivity between photons and electrons when it comes to quantum crystallographically relevant diffraction experiments.

Another point about the selectivity of QCBED is that the diffraction patterns are extremely sensitive to the degree of crystal perfection. The presence of crystal defects is often very strongly manifested by breaks in symmetry within the complex and detailed intensity distributions that make up convergent-beam electron diffraction (CBED) patterns. In practice, this means that should the experimenter observe breaks in symmetry within the diffraction patterns that they are collecting, these should be discarded and a new area of the specimen selected for CBED data collection. This is made simple by the ease with which an operator can switch between imaging the specimen in real space and imaging the diffraction pattern (at the push of a button).

A final point regarding the spatial selectivity of QCBED and the technique's future potential is made by Prof. Paul Midgley [9]:

"…highly accurate structure factors can be obtained from volumes of material much smaller than is possible even with synchrotron x-rays. This powerful combination offers the possibility of mapping bonding characteristics across heterogeneous samples (e.g., composites or quantum wells) that may not be achievable with any other method."

CBED – Some Background

Due to the strong interaction of electrons with matter, specimen thicknesses do not generally exceed about 200 nm (or 2,000 Å), which means that the shape transform of the specimen elongates the reciprocal lattice points into reciprocal lattice rods (often referred to as "*relrods*") in the direction of the specimen surface normal which is also usually in the same direction as the incident electron beam. In addition, the very short wavelengths ($\lambda \approx 0.02$ Å – 0.05 Å) of electrons used in TEMs means that the Ewald sphere is almost flat with respect to the reciprocal

lattice. What this means is that the Ewald sphere will intersect many *relrods*, especially if the incident beam is incident along a major zone axis in a crystal. Because of the proximity of many reflections to their respective Bragg conditions under these circumstances, there will be many reflections that have measurable and significant intensities in them.

If one then sets up an electron probe in the form of a convergent beam spanning a convergence angle of 2α, then the result is no longer a set of diffracted rays manifesting spots in a diffraction pattern but diffracted cones spanning the angle 2α, which manifest discs in the resulting CBED pattern. This is illustrated in figure 2.

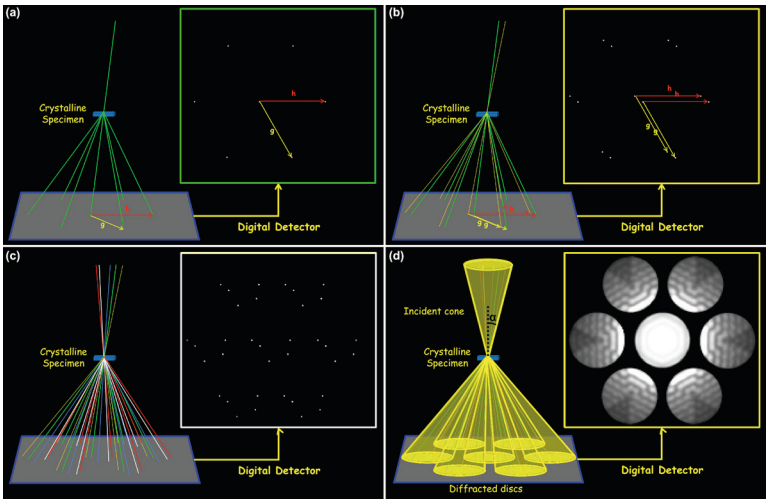


Figure 2. Schematic illustration of CBED and its relationship to parallel beam diffraction. **(a)** In parallel beam diffraction, the diffraction pattern is an arrangement of spots related by linear combinations of the two shortest non-colinear reciprocal lattice vectors (scattering vectors) appearing in the pattern, **g** and **h**. **(b)** If there are two beams incident upon the specimen in different directions, the result is two point patterns where the points in each individual pattern have the same geometry described by linear combinations of **g** and **h**. The two point-patterns are offset from one another by the difference in angle of incidence of the two incident beams. **(c)** By extension, the larger the number of incident beam directions, the more the diffraction pattern fills up with spots. **(d)** Taking this sequence to its extreme, the ultimate result is a cone of incident beams that results in diffracted discs where the position of a point in one disc is related to the same position in another disc by a scattering vector (reciprocal lattice vector) that is some linear combination of **g** and **h**. The cone convergence semi-angle is denoted by α.

The intensity within a diffracted disc as a function of angle is a 2-dimensional rocking curve and CBED patterns with detailed intensity distributions are extremely sensitive to the specimen thickness, the structure factors of the reflections whose rocking curves encompass or are close to their Bragg conditions and the structure factors of the scattering vectors that strongly couple different reflections in the dynamical scattering that dominates a particular diffraction geometry. In fact, diffraction geometry plays a large role in tailoring QCBED experiments and data collection to the measurement of specific structure factors and is covered in the next lecture – *QCBED Lecture 2: Experimental Procedures & Data Preparation.*

It follows from the previous paragraph that matching a calculated CBED pattern to an experimental one is a means for measuring a number of structure factors (those of the reflections at or near the Bragg condition plus those corresponding to the scattering vectors that strongly couple these reflections in dynamical scattering) and the specimen thickness. An example is given in figure 3, which illustrates the nature of QCBED pattern matching.
In the example shown for a CBED pattern obtained from corundum (α-Al_2O_3), five structure factors and the crystal thickness are refined as well as phenomenological absorption coefficients paired with each of the structure factors, the electron energy and the incident beam intensity. This results in 13 refined parameters. The diffraction pattern has been differentiated with respect to scattering angle to remove the diffuse background caused by inelastic scattering because the electron scattering calculations used in QCBED consider only elastic scattering (accounting for inelastic scattering would increase the length of the refinements by several orders of magnitude). After binning by a factor of 12 x 12 pixels, the number of independent data points matched is 4,032, which outnumbers the refined parameters more than 300 times. This leads to the very high precision associated with the results shown. QCBED refinements are usually very highly constrained by this rate of oversampling as well as the complexity of the rocking curve intensities within each reflection disc and the pattern as a whole.

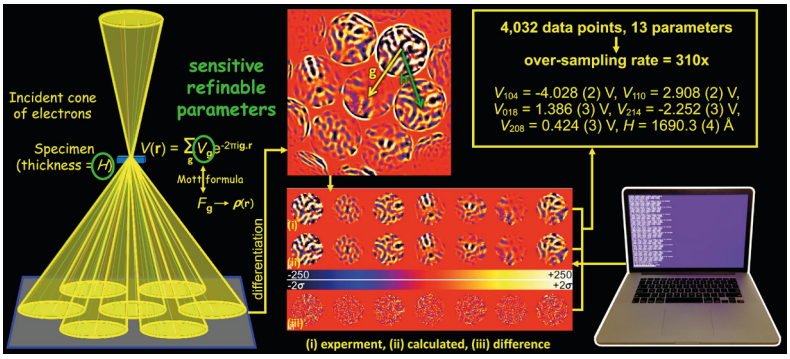


Figure 3. Schematic illustration of QCBED. A CBED pattern is collected using a digital detector and differentiated to remove the slowly varying diffuse background due to inelastically scattered electrons. The reflection discs to be pattern matched are extracted into a linear array which is input for the QCBED refinement algorithm. QCBED pattern matching involves the fitting of a computer-simulated CBED pattern to the experimental input by refining the parameters to which the intensities in the pattern are most sensitive. These include the structure factors of the reflections at or near the Bragg condition, the structure factors of the scattering vectors that couple these reflections in the highly dynamical scattering that occurs between them, and the specimen thickness. The matching process is heavily over-determined, as shown in the present example, leading to very high precision and accuracy of the measured structure factors. This example uses CBED data collected with 200 keV electrons at an incident beam orientation of 104 relative to [-441] in α-Al_2O_3 [10].

Conversion of structure factors from crystal potential to electron density via the Mott formula [2] further increases precision if $s = (\sin\theta)/\lambda$) is small (as in the case of the low order structure factors sensitive to bonding) because s^2 multiplies the crystal potential structure factors, V_g, and

therefore any uncertainty associated with them, in the process of determining F_g, the structure factors of the electron density (figure 1).

The accuracy of QCBED in measuring low order structure factors is further enhanced by the form of the atomic scattering factors for electrons, which increase very rapidly with decreasing s – more so than X-ray scattering factors as can be seen from the example of aluminium in figure 4. On the other hand, X-ray diffraction covers a much broader range of s in measuring structure factors and those of higher order (higher scattering angle) are much more reliably measured with X-rays than by QCBED. This is due to the difficulty of collecting CBED rocking curves spanning Bragg conditions for higher order reflections with sufficient two-dimensional detail to constrain QCBED refinements, as well as the lower sensitivity of electron scattering to core electron distributions as shown by the scattering factor plot in figure 4.

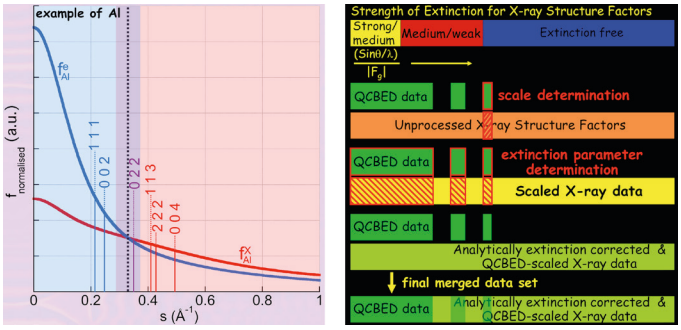


Figure 4. A comparison of electron and X-ray scattering factors (left) and a combination method for merging X-ray and QCBED-measured structure factor sets (right). If one considers the scattering factors for X-rays and electrons where the areas under both curves have been normalized with respect to one another, then it is evident that electrons are much more sensitive to the crystal potential than X-rays are to the electron density at low scattering angles ($s = (\sin\theta)/\lambda$). This region corresponds to the low-order, bonding-sensitive structure factors (shaded in blue). The region of s in which the two scattering factor curves cross over (at the vertical dotted line) is shaded in purple and beyond this (shaded in red), as s increases, X-rays become more sensitive. The present example shows the case for aluminium atoms and hkl corresponding to fcc elemental aluminium. The scattering factor curves correspond to the relativistic Hartree-Fock calculations of Doyle and Turner [11]. QCBED is known to be more accurate and precise when it comes to measuring low order structure factors [12 – 14], whilst X-ray diffraction is more accurate and precise when it comes to measuring the higher order structure factors. With this in mind and given that QCBED measurements are extinction-free and absolute (no scale factor is involved), the QCBED-measured structure factors can be used in combination with sets of structure factors measured by X-ray diffraction via a combination method such as the approach shown on the right of the figure (developed with and largely by Dr. V. Streltsov [15]). Such an approach produces a combined structure factor set that is much more useful than its individual components (QCBED and X-ray diffraction) in isolation.

So, in summary, QCBED is very effective at measuring a small number (a few tens) of low to medium order structure factors with high accuracy and precision whilst X-ray diffraction can measure many more structure factors (thousands) covering a wide range of s from low to high order of hkl. The lower order and strong structure factors measured by X-ray diffraction are highly susceptible to extinction (a consequence of multiple scattering) [5 – 8] and the scale of the X-ray diffraction measured structure factors must be determined using an external calibrant

(often just the independent atom model values for higher order structure factors) [8]. The availability of absolute and extinction-free structure factors from QCBED thus presents an opportunity to merge them with any set of X-ray diffraction-measured structure factors for the same material. A combination method for merging QCBED and X-ray diffraction structure factors sets, developed by Dr. V. Streltsov [15] in conjunction with this author, is shown in figure 4. This approach takes advantage of the large range and number of structure factors that can be measured relatively quickly with X-ray diffraction experiments by merging them with QCBED-measured structure factors that have a much smaller range, take significantly longer to determine, but can be used to determine the scale of the X-ray structure factors and correct them for extinction.

Considering how the diffracted intensities in computer-simulated CBED patterns are calculated, is instructive with respect to considering how QCBED could be adapted to future quantum crystallography applications. Two completely different formalisms exist for a complete description of dynamical scattering of electrons: (i) the Bloch-wave method [16, 17] and (ii) the multislice [18] formulation. The fundamental components of both are illustrated in figure 5 with details in the caption.

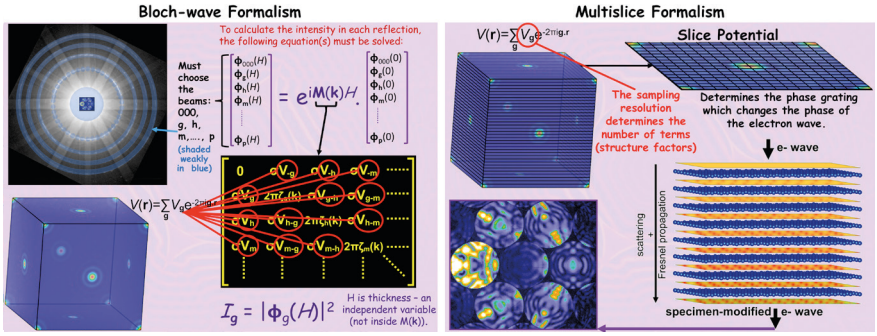


Figure 5. Two independent formalisms for calculating dynamical electron scattering – the Bloch-wave [16, 17] and multislice formalisms [18]. The Bloch-wave method formulates dynamical scattering as a set of N simultaneous equations where N is the number of reflections that are chosen for inclusion in the scattering equations. These are reflections within a user-specifiable range of the Bragg condition and in the illustration above, these appear as reflections with significant intensities in a diffraction pattern that includes higher angle reflections. Diagonalization of the scattering matrix results in the solution of the simultaneous dynamical scattering equations and yields the intensities in all reflections included in the calculation. The solution process in this case is thickness independent and specimen thickness is introduced after the solution of the scattering equations as an independent variable. In contrast, the multislice formalism slices the material into layers of crystal potential that are periodic in the plane of the slices which are perpendicular to the incident beam direction (there is no requirement for periodicity in the beam direction). The projected potential within each slice determines a phase grating which modifies the electron wave. The resolution with which this projected potential (and therefore the phase grating) is computed determines the number of structure factors (and scattering vectors) included in the simulation of diffracted intensities. The electron wave is Fresnel-propagated between slices and the entire calculation is a series of convolutions of scattering and Fresnel propagation until the exit face of the specimen has been reached. The specimen thickness and slice thicknesses determine the total number of slices and thus, the number of scattering and propagation convolutions that are applied in a calculation. As a result, these types of calculations cannot treat the thickness as a separate variable as it is a fundamental component of the calculation process.

It is worth remembering that whilst QCBED matches only a few reflections (see figure 3 for an example), there are many others that contain significant intensity (see figure 5) that must be modelled correctly because of their collective effects on the reflections being pattern matched. In the Bloch-wave formalism, these reflections must be identified individually, usually by way of a reflection selection algorithm with user-specified selection criteria. In the multislice method, the number and indices of reflections are set by the slice sampling resolution.

The two different approaches to calculating CBED patterns leads to different requirements, structural constraints, computational expenses and modes of applicability when it comes to QCBED. These are summarized in table 2.

Something that stands out from table 2 is the difference in the rate of change of the computation time as a function of the number of reflections included in each type of calculation. The multislice method is also linearly dependent on the number of slices used in the calculation (the thicker the probed region of the specimen and the thinner the slices used, the more convolutions are required), but changes at a much slower rate in terms of the number of reflections included than the Bloch-wave approach. It is, however, worth realizing that for small molecule crystallography and fully converged calculations of intensities by both methods, the calculation times are roughly the same. A typical QCBED refinement like that in the example of figure 3 would take about 24 to 48 hours incorporating 256 reflections into both the Bloch-wave and multislice methods for a specimen thickness of approximately 1,500 Å. It is only when unit cells become larger and the atomic motifs more complex that the multislice method will start to become more efficient than the Bloch-wave approach.

Table 2. A comparison of the Bloch-wave and multislice formalisms in their application to QCBED.

QCBED with Bloch-waves	**QCBED with multislice**
• Specimen thickness is extrinsic to the solution of the scattering equations.	• Specimen thickness is intrinsic to the electron scattering calculation.
• The probed volume of the specimen must have 3-dimensional periodicity (the stacked Bloch-wave approach [19] is a way of breaking down this requirement but has limitations).	• The probed volume of the specimen need only have 2-dimensional periodicity in the plane of each slice, ***making it possible to refine atomic structure and structure factors in layered nanocomposites.***
• Reflection selection for solving the scattering equations is on an individual basis and completely user-specifiable.	• Reflection selection is collective and dependent on the slice sampling resolution
• Easy to compute diffraction patterns for any orientation of the incident electron beam.	• The slicing direction limits the range of orientations for which diffraction patterns can be calculated.
• The calculation of diffracted intensities is based on the diagonalization of an N x N matrix where N is the number of reflections chosen for inclusion in the calculation.	• The calculation of diffracted intensities is based on a series of fast Fourier transforms (FFTs). The number of reflections is $N = 2^p \times 2^r$, where p and r are integers and 2^p and 2^r are dimensions of the mesh used for sampling the crystal potential in each slice.
• Computation times are proportional to N^3.	• Computation times are proportional to $N.\log(N).H/\Delta H$, where H is the specimen thickness and ΔH is the mean thickness of each slice (note that the slices do not need to have the same thicknesses)

QCBED Development and Quantum Crystallography

A final point is the adaptability of QCBED to new directions in quantum crystallography. Figure 6 illustrates two different semantics for the refinement of electron distribution (or crystal potential) that are fully commutative in QCBED (and indeed in X-ray diffraction analyses). The first shows the way in which QCBED has always been performed, that is: starting with an initial set of structure factors, usually from an independent atom model (IAM) or possibly even a solid-state theory that includes bonding effects like density functional theory (DFT), which are then refined until the best fit between the calculated and experimental CBED intensities is reached. The refined structure factors are then used to compute the 3-dimensional crystal potential or electron density.

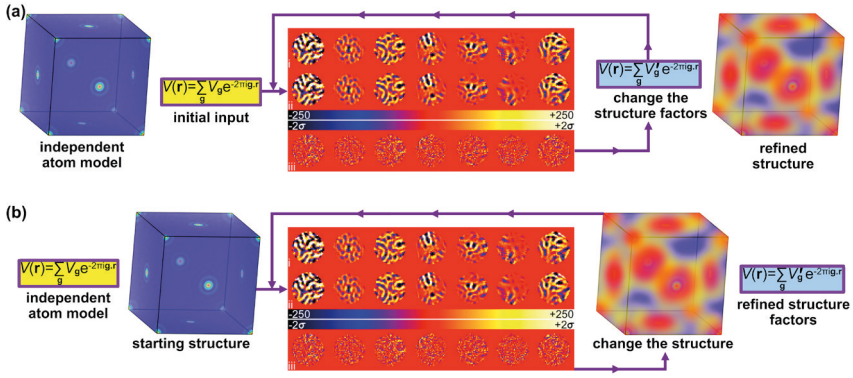


Figure 6. Two philosophies for QCBED. **(a)** All QCBED is currently practiced in this form, where structure factors are refined from starting values (usually obtained from an IAM) to optimized values by minimizing the mismatch between the calculated and experimental CBED intensity distributions. The variable parameters are the structure factors and the refined structure is the consequence of the refinement of the structure factors. This process is carried out in reciprocal space with the outcome in real space only being considered AFTER the refinement process. **(b)** A conceptually different approach would be to base the refinement in real space by changing the electron distribution and even atomic positions (both being real space parameters). The refinement of structure factors then becomes a consequence of the real space structure refinement. Whilst the equivalence of **(a)** and **(b)** is obvious the differences that result from their practical implementation from a computational and materials modelling perspective open up new possibilities.

An alternative to this is shown in figure 6(b), where the object of the refinement is the 3-dimensional crystal potential or electron distribution rather than the structure factors themselves. Here the by-products of the refinement are the structure factors. This has many ramifications including the ability to use other methods for generating potential or electron distributions like multipoles, DFT and other structural models as "plug-ins". This would then allow parameters that control these "plug-ins" to be refined instead. A single parameter in any one of these structural models may affect a large number of structure factors simultaneously but instead of having many extra refinement parameters corresponding to each affected structure factor, there would be just one or two model parameters models as refinement variables. This

would provide an efficient means of directing solid state theory using experimental data, as some have already doing for quite some time in X-ray diffraction [20].

These notes are supplementary to the lecture content.

References

[1] L. Massa, L. Huang, J. Karle, *Int. J. Quant. Chem. Quant. Chem. Symp.* **29** (1995), 371 – 384.

[2] N.F. Mott, H.S.W. Massey, The theory of atomic collisions. In: *International Series of Monographs on Physics, 3rd Ed.* Clarendon Press, Oxford (1965).

[3] P.N.H. Nakashima, *Struct. Chem.* **28** (2017), 1319 – 1332.

[4] A. Genoni, L. Bučinsky, N. Claiser, J. Contreras-Garcia, B. Dittrich, P.M. Dominiak, E. Espinosa, C. Gatti, P. Giannozzi, J.-M. Gillet, D. Jayatilaka, P. Macchi, A.Ø. Madsen, L.J. Massa, C.F. Matta, K.M. Merz, P.N.H. Nakashima, H. Ott, U. Ryde, K. Schwarz, M. Sierka, S. Grabowsky, *Chem. Eur. J.* (2018). *In press.*

[5] W.H. Zachariasen, *Acta Cryst.* **16** (1963), 1139 – 1144.

[6] P.J. Becker, P. Coppens, *Acta Cryst. A* **30** (1974), 129 – 147.

[7] E.N. Maslen, N. Spadaccini, *Acta Cryst. A* **49** (1993), 661 – 667.

[8] P. Coppens, *X-ray Charge Densities and Chemical Bonding* (International Union of Crystallography, Oxford University Press, New York, 1997).

[9] P.A. Midgley, *Science* **331** (2011), 1528 – 1529.

[10] P.N.H. Nakashima, B.C. Muddle, *Phys. Rev. B* **81** (2010), 115135.

[11] P.A. Doyle, P.S. Turner, *Acta Cryst. A* **24** (1968), 390 – 397.

[12] J.M. Zuo, *Rep. Prog. Phys.* **67** (2004), 2053 – 2103.

[13] J.C.H. Spence, *Acta Cryst. A* **49** (1993), 231-260.

[14] J.M. Zuo, J.C.H. Spence, *Advanced Transmission Electron Microscopy* (Springer, New York, 2017).

[15] V.A. Streltsov, P.N.H. Nakashima, A.W.S. Johnson, *Microsc. Microanal.* **9** (2003), 419 – 427.

[16] H. Bethe, *Ann. Phys.* **392** (1928), 55 – 129.

[17] F. Bloch, *Z. Physik* **52** (1929), 555 – 600.

[18] J.M. Cowley, A.F. Moodie, *Acta Cryst.* **10** (1957), 609 – 619.

[19] R.S. Pennington, F. Wang, C.T. Koch, *Ultramicroscopy* **141** (2014), 32 – 37.

[20] D. Jayatilaka, D.J. Grimwood, *Acta Cryst. A* **57** (2001), 76 – 86.

Recommended Reading

Early QCBED (before electron energy filters, digital detectors and powerful computers)

[21] C.H. MacGillavry, *Nature* **145** (1940), 189 – 190.

[22] C.H. MacGillavry, *Physica* **7** (1940), 329 – 343.

[23] P. Goodman, G. Lehmpfuhl, *Acta Cryst.* **22** (1967), 14 – 24. (*Multislice*)

[24] R. Voss, G. Lehmpfuhl, P.J. Smith, *Z. Naturforsch. A* **35** (1980), 973 – 984. (*Multislice*)

QCBED Revival (advent of electron energy filters, digital detectors and rising computing power)

One-Dimensional (Systematic Row) QCBED (Bloch-wave)

[25] J.M. Zuo, J.C.H. Spence, M. O'Keeffe, *Phys. Rev. Lett.* **61** (1988), 353 – 356.

[26] J.C.H. Spence, J.M. Zuo, *Electron Microdiffraction* (Plenum Press, New York, 1992).

[27] J.M. Zuo, *Acta Cryst. A* **49** (1993), 429 – 435.

[28] C. Deininger, G. Necker, J. Mayer, *Ultramicroscopy* **54** (1994), 15 – 30.

[29] R. Holmestad, J.M. Zuo, J.C.H. Spence, R. Høier, Z. Horita, *Philos. Mag. A* **72** (1995), 579 – 601.

[30] J.M. Zuo, M. Kim, M. O'Keeffe, J.C.H. Spence, *Nature* **401** (1999), 49 – 52.

Two-Dimensional QCBED (Bloch-wave)

[31] D.M. Bird, M. Saunders, *Ultramicroscopy* **45** (1992), 241 – 251. **(*Zone axis*)**

[32] M. Saunders, D.M. Bird, N.J. Zaluzec, W.G. Burgess. A.R. Preston, C.J. Humphreys, *Ultramicroscopy* **60** (1995), 311 – 323. **(*Zone axis*)**

[33] M. Saunders, A.G. Fox, P.A. Midgley, *Acta Cryst. A* **55** (1999), 471 – 479. **(*Zone axis*)**

[34] K. Tsuda, M. Tanaka, *Acta Cryst. A* **55** (1999), 939 – 954. **(*Whole pattern*)**

[35] V.A. Streltsov, P.N.H. Nakashima, A.W.S. Johnson, *J. Phys. Chem. Solids* **62** (2001), 2109 – 2117. **(*Near zone axis*)**

QCBED Modern Era

One-Dimensional (Systematic Row) QCBED (Bloch-wave)

[36] B. Jiang, J.M. Zuo, N. Jiang, M. O'Keeffe, J.C.H. Spence, *Acta Cryst. A* **59** (2003), 341 – 350.

[37] J. Friis, G.K.H. Madsen, F.K. Larsen, B. Jiang, K. Marthinsen, R. Holmestad, *J. Chem. Phys.* **119** (2003), 11359 – 11366.

[38] J. Friis, B. Jiang, K. Marthinsen, R. Holmestad, *Acta Cryst. A* **61** (2005), 223 – 230.

[39] R. Sæterli, E. Flage-Larsen, J. Friis, O.M. Løvvik, J. Pacaud, K. Marthinsen, R. Holmestad, *Ultramicroscopy* **111** (2011), 847 – 853.

Two-Dimensional QCBED

[40] K. Tsuda, Y. Ogata, K. Takagi, T. Hashimoto, M. Tanaka, *Acta Cryst. A* **58** (2002), 514 – 525. **(*Bloch-wave, Whole pattern*)**

[41] V.A. Streltsov, P.N.H. Nakashima, A.W.S. Johnson, *Microsc. Microanal.* **9** (2003), 419 – 427. **(*Multislice & Bloch-wave, Near zone axis*)**

[42] Y. Ogata, K. Tsuda, Y. Akishige, M. Tanaka, *Acta Cryst. A* **60** (2004), 525 – 531. **(*Bloch-wave, Whole pattern*)**

[43] P.N.H. Nakashima, *J. Appl. Cryst.* **38** (2005), 374 – 376. **(*Bloch-wave, Near zone axis*)**

[44] K. Tsuda, D. Morikawa, Y. Watanabe, S. Ohtani, T. Arima, *Phys. Rev. B* **81** (2010), 180102(R). **(*Bloch-wave, Whole pattern*)**

[45] X. Sang, A. Kulovits, J.M.K. Wiezorek, *Acta Cryst. A* **66** (2010), 694 – 702. **(*Bloch-wave, Near zone axis*)**

[46] X. Sang, A. Kulovits, J.M.K. Wiezorek, *Acta Cryst. A* **67** (2011), 229 – 239. **(*Bloch-wave, Near zone axis*)**

Differential QCBED

[47] P.N.H. Nakashima, *Phys. Rev. Lett.* **99** (2007), 125506. **(*Bloch-wave, Near zone axis*)**

[48] P.N.H. Nakashima, B.C. Muddle, *Phys. Rev. B* **81** (2010), 115135. **(*Bloch-wave, Near zone axis*)**

[49] P.N.H. Nakashima, A.E. Smith, J. Etheridge, B.C. Muddle, *Science* **331** (2011), 1583 – 1586. **(*Bloch-wave, Near zone axis*)**

[50] D. Peng, P.N.H. Nakashima, *J. Appl. Cryst.* **50** (2017), 602 – 611. **(*Multislice, Near zone axis*)**

Periodic systems: models and strategies

Bartolomeo Civalleri
Department of Chemistry, University of Turin, Turin, Italy
bartolomeo.civalleri@unito.it

The purpose of these notes is to provide a concise description of the role of translational symmetry in dealing with extended systems. This reflects on the way a model system (i.e. a periodic structure) can be constructed and more importantly on how its electronic structure (i.e. crystalline orbitals) is affected. In doing that we will refer to crystalline orbitals as a linear combination of Bloch functions expressed in terms of atomic orbitals (i.e. atom-centred Gaussian-type functions). These notes are meant to serve as a basic introduction. All topics are covered in more detail in the references given at the end, in particular see ref. [1] and [2].

Introduction

Extended systems exhibit a very large and complex structure in which phenomena are not localized but depend on the whole structure. In many cases, they have been considered as synonymous of solids, low-dimensionality systems and nanostructures. Here we will consider *spatially* extended systems in that they are extended in some way throughout space.
Among them, periodic systems are a special class of extended systems in which translational symmetry plays a crucial role. In fact, they can be considered as *translationally invariant* extended systems. It turns out that the system under study is an infinite and perfect object that can be periodic in 1-, 2- and 3-directions (See Figure 1).

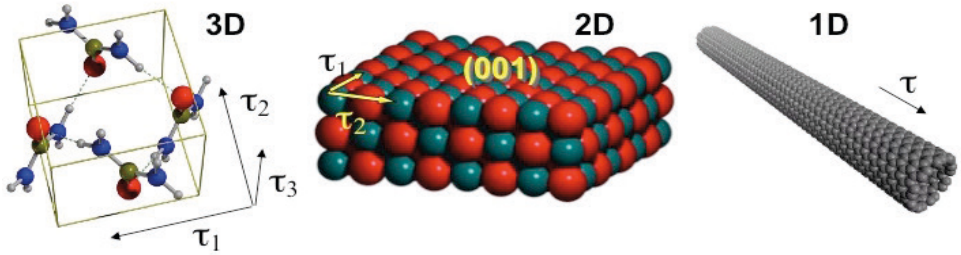


Figure 1. Example of periodic systems of different dimensionality with translational invariance in 3-, 2- and 1-directions. Translation vectors are indicated as τ_i.

How can periodic systems be modelled? Can we adopt the same theoretical methods (e.g. ab initio) as for molecules? Do we need special tricks?
Translation invariance has a series of interesting properties with important consequences on simplification of the problem and the implementation of efficient algorithms [1,2].

Periodic model systems

When modelling a material, usually in the solid state, it is not always possible to treat it as it is (as a whole) because of its large size (e.g. zeolites, MOFs) or its complex structure (e.g. amorphous solids, low-dimensionality systems, nanostructures). Therefore, in many cases, the real system must be mimicked through a *model system*. To do that, one has to consider that solids can show different combinations of structure topology and chemical bonding. For instance, the structure of crystalline materials can be dense, porous (e.g. nano/micro/meso) or layered with chemical bonding ranging from ionic to covalent (with intermediate semi-ionic or semi-covalent character) and from metallic to molecular. 3D structures become even more complicate when dealing with disordered materials, amorphous solids or defective systems. Furthermore, the realm of nanoworld has recently attracted a lot of interest in materials science. In that case, nanostructures such as nanorods, nanowires and nanotubes have become more and more investigated.

Within a periodic approach, the real system is then modelled as a *periodic model system* in which a repeat unit (hereafter denoted as *unit cell*) is replicated along periodic (infinite) directions. Periodicity can then be exploited in one- (1D), two- (2D) and three- (3D) directions. For 1D and 2D, the system remains of finite size along the other directions. Accordingly, periodic model systems can be regarded as:

- *3D model*. This is the natural choice for inherently crystalline materials. However, solids are never perfect like in defective systems (both extended and local defects) and disordered (amorphous) solids. In such cases translational symmetry is apparently lost. Nevertheless, the periodic approach can be recovered through the use of a *supercell model* [2,3] in which the replicated unit cell is large enough to include the imperfection (e.g. local defect, disorder, …). The size and shape of the supercell depend on the examined systems. For point defects, defect-defect interactions should be avoided among periodic replicas. In disordered/amorphous systems, the supercell must be large enough to guarantee only short-range order and amorphization of the structure. The supercell model can also be adopted to study heterostructures.
- *2D model*. An ideal crystal is an infinite object while a real crystal is limited by different surfaces. Surfaces are important in many chemical and physical phenomena as catalysis, adsorption, corrosion and oxidation. Yet, surfaces show a rich "chemical complexity" because they can be internal (e.g. porous materials) and external (e.g. surfaces of metals and oxides grains). In addition, steps, kinks, and defects can also be present. In this case, the periodic approach is exploited by using the so-called *slab model* [2-4]. A thin film is cut from the solid, comprised by a few atomic layers, parallel to a given *(hkl)* face. The structure remains periodic along two directions and finite in the perpendicular one. Surface properties should converge with the slab thickness. Note that not all crystalline surfaces are physically stable or worthy of investigation.

- *1D model*. A periodic approach can be easily employed to model polymers with regular or helicoidal structure. In the realm of nanostructures, nanowires and nanotubes can be modelled as one-dimensional structures. For instance, a nanorod can be created by cutting a 2D model parallel to another *(h'k'l')* face [5] or a nanotube can be generated by wrapping an initial 2D model according to a rolling vector [6]. The latter defines the diameter and chirality of the tube.
- *0D model*. This can be considered the limit case of a periodic approach in which translational symmetry is completely lost, but point symmetry still remains. In materials modelling this is known as a *cluster model*. It can be used to model nanoparticles [5].

In the next section, we will discuss in more detail how translational symmetry can be exploited in a 3D model [2,8]. Results will be completely general and can be extended to lower dimensionality models by removing translational symmetry along one or two directions. It is assumed that basic knowledge of Crystallography is known.

Direct and reciprocal lattices

A 3D structure is a natural model of a crystalline system. It can then be described as an ordered arrangement of atoms or molecules in three dimensions []. The set of points at which atoms (or molecules) are repeated throughout space by translational symmetry forms the *direct lattice*. By definition, a lattice is translationally invariant. Points are repeated at intervals of length a_1, a_2 and a_3 along three non-coplanar directions. The three constants a_1, a_2 and a_3 are called lattice parameters, and the vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$, oriented in the same three non-coplanar directions with the lattice parameters are the basis vectors. Basis vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ define the unit cell. Lattice parameters and angles between the basis vectors are called cell parameters.

A general lattice vector can always be expressed as a linear combination of the basis vectors:

$$\boldsymbol{g} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$$

where coefficients n_i are integers.

Note that translations along lattice vectors are symmetry operations.

The unit cell corresponds to a volume in space that fills space entirely when translated by all lattice vectors. The choice of the unit cell is not unique. A set of parallelepipeds defined by the basis vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$, with the $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ as much orthogonal as possible and the cell as symmetric as possible can be classified in 14 types denoted Bravais lattices. A unit cell containing only one lattice point is called primitive cell.

The position **r** of an atom in the unit cell is usually expressed in terms of fractional coordinates f_1, f_2, f_3 such that

$$\boldsymbol{r} = f_1\boldsymbol{a}_1 + f_2\boldsymbol{a}_2 + f_3\boldsymbol{a}_3$$

f_1, f_2 and f_3 values are in the interval 0 to 1.

In the interpretation of many physical properties of crystals (e.g. diffraction, electronic structure, phonons…) it is very useful to introduce the so-called *reciprocal lattice*. Any direct lattice can be associated to a new lattice defined through reciprocal lattice basis vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ as:

$\boldsymbol{a}_i \cdot \boldsymbol{b}_j = 2\ _{ij}$

This is an orthogonality condition in which $_{ij}$ is the Kronecker delta.
Equivalently:

$\boldsymbol{b}_1 = 2\ / \mathrm{V}\ \boldsymbol{a}_2 \times \boldsymbol{a}_3\ ;\ \boldsymbol{b}_2 = 2\ / \mathrm{V}\ \boldsymbol{a}_3 \times \boldsymbol{a}_1\ ;\ \boldsymbol{b}_3 = 2\ / \mathrm{V}\ \boldsymbol{a}_1 \times \boldsymbol{a}_2$

where $\mathrm{V} = \boldsymbol{a}_1 \cdot \boldsymbol{a}_2 \times \boldsymbol{a}_3$ and $\mathrm{V}^* = (2\ )^3 / \mathrm{V}$ are the volumes of the unit cell in the direct and reciprocal space, respectively.
Any reciprocal lattice vector can be expressed as a linear combination of the basis vectors with integer coefficients (m_i) such as

$\mathbf{K} = m_1\boldsymbol{b}_1 + m_2\boldsymbol{b}_2 + m_3\boldsymbol{b}_3$

A general position in the reciprocal lattice is a linear combination of basis vectors with real coefficients:

$\mathbf{k} = k_1\boldsymbol{b}_1 + k_2\boldsymbol{b}_2 + k_3\boldsymbol{b}_3$

A useful way to define a primitive unit cell in the reciprocal space is through the Dirichlet construction. By connecting one reciprocal lattice point to all its nearest neighbors and letting orthogonal planes pass through their midpoints, a region of space which is closer to one lattice point than to anyone else is delimited. This corresponds to the Wigner-Seitz cell in the reciprocal space, which is also known as the first Brillouin Zone. There are 14 different types of first Brillouin Zones corresponding to the 14 reciprocal Bravais lattices. See ref. [7] for more details.

Bloch's theorem, Bloch functions and crystalline orbitals

When solving the Schrödinger equation for a 3D periodic model system, one can think to treat it as a huge molecule [2,8]. It is not unexpected, by intuition, that the calculation of the electronic structure and related properties rapidly becomes an unattainable problem.
In fact, in order to calculate molecular structures and properties, it is necessary to determine the eigenfunctions Ψ and eigenvalues E of the Schrödinger equation

$\mathrm{H}\Psi = \mathrm{E}\Psi$

where H is the electronic Hamiltonian. For systems of interest in chemistry, one normally resorts to the variational approach. The most common way to implement this approach is to express the wave function in terms of one-electron spin-orbitals $\{\phi\}$ (molecular orbitals, MO) that in turn are written as a linear combination (LC) of known basis functions $\{\chi_j\}$, usually denoted as *basis set*:

$\phi(\mathbf{r}) = \Sigma_j c_j \chi_j(\mathbf{r})$

When the basis set is built with atomic orbitals (AO), this corresponds to the well-known MO-LCAO approximation.
One-electron spin-orbitals are then combined in an anti-symmetrized wave function known as Slater determinant, which is commonly adopted in the frame of the Hartree-Fock method and in the Kohn-Sham formalism of the Density Functional Theory (see Giannozzi's notes).
When the basis functions are orthonormal, this leads to the solution of an eigenvalue problem that can be written in matrix notation as

HC=CE

Where **H** is the Hamiltonian matrix in terms of the basis functions, **C** is the matrix of the coefficients and **E** is the eigenvalue matrix. The size of the matrices depends on the number of basis functions. It is then evident now that if an extended system is treated as a huge molecule, one has to deal with infinitely sized matrices. Fortunately, the exploitation of translation symmetry in periodic model systems makes this computational problem solvable.

Bloch's theorem and Bloch functions

Let us consider a Hamiltonian operator: H = T + V(**r**), which contains a periodic potential, such that

$V(\mathbf{r}) = V(\mathbf{r}+\mathbf{g})$,

For any **g** then the wavefunction exhibits the following translational symmetry:

$\Phi(\mathbf{r}+\mathbf{g}) = e^{i\mathbf{k}\cdot\mathbf{g}}\,\Phi(\mathbf{r})$,

This is the main consequence of the *Bloch's theorem*. Thus, the wave function at translationally equivalent points of the direct lattice corresponds to the initial wave function at **r** times a phase factor $e^{i\mathbf{k}\cdot\mathbf{g}}$ known as the *Bloch phase factor* in which **k** is a *wave vector*.

If $\mathbf{k}' = \mathbf{k} + \mathbf{K}$, where **K** is a reciprocal lattice vector, the Bloch phase factor becomes

$e^{i\mathbf{k}'\cdot\mathbf{g}} = e^{i\mathbf{k}\cdot\mathbf{g}}\, e^{i\mathbf{K}\cdot\mathbf{g}} = e^{i\mathbf{k}\cdot\mathbf{g}}$

in which it has been used that $e^{i\mathbf{K}\cdot\mathbf{g}} = e^{2\pi i\, mj\, nj} = 1$ for the orthonormality condition between reciprocal and direct lattice vectors. This means that **k** can be limited to the first Brillouin zone. Each function that satisfies the Bloch theorem is an eigenfunction of the lattice translation operators and is called *Bloch Function*.
Alternatively, by using group theory, it can be shown that $e^{i\mathbf{k}\cdot\mathbf{g}}$ is the eigenvalue of the translation operator T which commutes with the Hamiltonian H: T H=H T (because of the crystal symmetry) and shares the same eigenfunctions:

$$T\,\Phi(\mathbf{r}) = \Phi(\mathbf{r}+\mathbf{g}) = e^{i\mathbf{k}\cdot\mathbf{g}}\,\Phi(\mathbf{r})$$

Each eigenfunction of H can then be labelled with **k** indicating the irreducible representation of the translation group: $\Phi^{\mathbf{k}}(\mathbf{r})$ (or equivalently $\Phi(\mathbf{k};\mathbf{r})$)

Periodic boundary conditions (Born-von Karman)

The real crystal is macroscopic but finite. It can be viewed as a parallelepiped containing $N = N_1$ x N_2 x N_3 unit cells with sides N_i $\mathbf{a}_i$ (i=1,3).

When N_i grows to infinity, the ratio of the number of atoms at the surface to the total number of atoms in the crystal goes to zero. It turns out that a macroscopic crystal mostly exhibits properties and features of the bulk material, so that bulk properties should be insensitive to the surface and to the boundary conditions imposed on the wave function.

To assure periodicity the *Born-von Karman periodic boundary conditions* can be adopted:

$$\Phi^{\mathbf{k}}(\mathbf{r}+N_j\boldsymbol{a}_j) = \Phi^{\mathbf{k}}(\mathbf{r})$$

as if the crystal was a three-dimensional infinite array of identical and contiguous finite crystals. The dependence of each Φ from **k** has been highlighted.

By applying the Bloch theorem one gets:

$$\Phi^{\mathbf{k}}(\mathbf{r}+N_j\boldsymbol{a}_j) = e^{i\,Nj\,\mathbf{k}\cdot\mathbf{aj}}\,\Phi^{\mathbf{k}}(\mathbf{r}) = \Phi^{\mathbf{k}}(\mathbf{r})$$

which implies that the phase factor is equal to one: $e^{i\,Nj\,\mathbf{k}\cdot\mathbf{aj}} = 1$.

But if

$$e^{i\,Nj\,\mathbf{k}\cdot\mathbf{aj}} = e^{i\,Nj\,kj\,\mathbf{bj}\cdot\mathbf{aj}} = e^{2\pi i\,Nj\,kj} = 1$$

this equality is fulfilled by constraining k_j (real coefficients) to be

$$k_j = n_j / N_j$$

with n_j being an integer.

There exist then N_1 x N_2 x N_3 **k** points in every reciprocal lattice cell, each of which can be written by the reciprocal lattice basis vectors as

$$\mathbf{k} = \frac{n_1}{N_1}\mathbf{b}_1 + \frac{n_2}{N_2}\mathbf{b}_2 + \frac{n_3}{N_3}\mathbf{b}_3$$

If n_j is such that $0 \le n_j \le N_j$ for every j **k** belongs to the origin cell of the reciprocal lattice. When N_j tends to infinity, **k** points are very close one another and at the limit they completely fill the space. It turns out that **k** can be considered as a continuous variable.

Since **k** points can be limited to the first Brillouin Zone, as shown above, the sum over **k**-points can then be transformed into integrals over the first Brillouin Zone.

Notably, the use of Bloch functions allows us to take advantage of their properties to simplify the solution of the Schrödinger equation for a periodic model system. In fact, by referring to a basis set of n_{BF} Bloch functions, the Hamiltonian matrix (**H**) becomes block-diagonal (see Figure 2), with each block referring to one particular point **k** in the reciprocal space. The solution of the Schrödinger equation for a periodic system (of infinite size) is then transformed into an infinite set of eigenvalue equations of finite size (i.e. n_{BF} x n_{BF}).

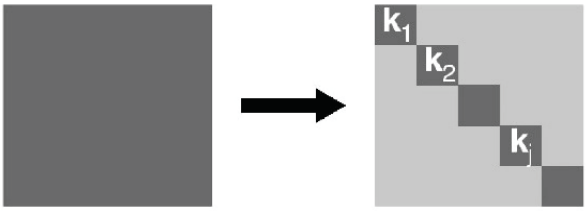


Figure 2. Factorization of the Hamiltonian matrix in a block-diagonal form when using Bloch functions.

Apparently, this is not a great advantage but owing to the usually smooth change of the eigenvalues and the eigenvectors with **k**, it is generally possible to sample matrix **H** at a finite number of points.
Therefore, the Schrödinger equation in a periodic approach is solved at different points in the first Brillouin zone:

$$H\Psi_n(\mathbf{k};\mathbf{r})=E_n(\mathbf{k})\Psi_n(\mathbf{k};\mathbf{r})$$

The eigenfunctions of the equation above are the *crystalline orbitals (CO)*. Eigenvalues $E_n(\mathbf{k})$ are characterized by two indices: **k** and *n*. As discussed above, **k** represents a point in the reciprocal space belonging to the first Brillouin zone. Since **k** is a continuous variable there is a *band of eigenvalues* tagged by the band index *n.* This leads to a spectral distribution of the electronic levels that represents the electronic band structure of the crystal. As an example, the band structure of silicon (semiconductor) and beryllium (metal) are shown in Figure 3.

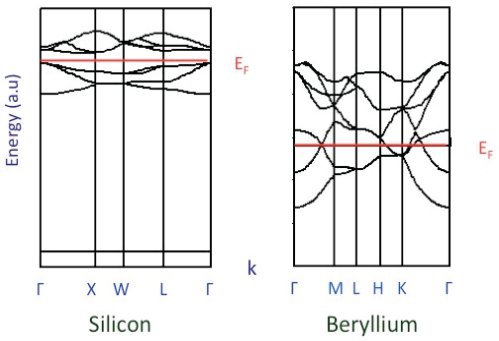


Figure 3. Band structure of silicon and beryllium. The Fermi energy (E_F) is highlighted in red.

Band structure of solids can be interpreted by using the language of band theory: energy gaps, band width, valence bands, conduction bands, Fermi energy, Fermi surface… and related quantities like the density of states (DOS). However, this is not in the scope of the present notes, see for instance ref. [9]

Localized basis functions

Crystalline orbitals are then linear combination of Bloch functions:

$\Psi_n(\mathbf{k};\mathbf{r}) = \Sigma_j\, c_{j,n}(\mathbf{k})\, \Phi_j(\mathbf{k};\mathbf{r})$

They are usually expressed in terms of either plane waves or localized atomic orbitals. Hereafter we refer to localized atom-centred functions

$\Phi_j(\mathbf{k};\mathbf{r}) = N^{-½} \Sigma_g\, e^{i\mathbf{k}\cdot\mathbf{g}} H_j(\mathbf{r} - \mathbf{r}_j - \mathbf{g})$

in that atomic orbitals (AOs) are centred at the position $\mathbf{r}_j$ located in the cell corresponding to the direct lattice vector $\mathbf{g}$. Usually, $\mathbf{r}_j$ corresponds to the atomic positions in the unit cell. It can be easily showed that $\Phi_j(\mathbf{k};\mathbf{r})$ satisfies the Bloch's theorem. AOs are commonly linear combinations of the products of *Gaussian-type functions* by real, solid harmonics [1,2]. In analogy with the well-known MO-LCAO, here we refer to it as the CO-LCAO approximation. Coefficients $c_{j,n}(\mathbf{k})$ are obtained within a variational approach by solving a matrix equation at different $\mathbf{k}$-points:

$\mathbf{H}(\mathbf{k})\mathbf{C}(\mathbf{k}) = \mathbf{S}(\mathbf{k})\mathbf{C}(\mathbf{k})\mathbf{E}(\mathbf{k})$

where $\mathbf{S}(\mathbf{k})$ is the overlap matrix between Bloch functions, $\mathbf{C}(\mathbf{k})$ is the matrix of the coefficients, $\mathbf{E}(\mathbf{k})$ is the eigenvalue matrix, $\mathbf{H}(\mathbf{k})$ is the Fock/Kohn-Sham matrix in the reciprocal space. With the orthonormality condition at every $\mathbf{k}$ point: $\mathbf{C}(\mathbf{k})^{\dagger}\mathbf{S}(\mathbf{k})\mathbf{C}(\mathbf{k}) = \mathrm{I}$.

The first Brillouin zone can be sampled on a finite set of $\mathbf{k}$-points. The number of $\mathbf{k}$ points to be considered is usually relatively small and solving the Schrödinger equation in the reciprocal space is a feasible method.

$\mathbf{H}(\mathbf{k})$ is obtained as the Fourier transform of the Fock/Kohn-Sham matrix in the direct space $\mathbf{H}(\mathbf{g})$. In terms of a localized basis set, the element of the Hamiltonian matrix in the direct space is computed as

$H_{ij}(\mathbf{g}) = \langle \chi_i(\mathbf{r} - \mathbf{r}_i - \mathbf{g}') | H | \chi_j(\mathbf{r} - \mathbf{r}_j - \mathbf{g}'') \rangle$

where the direct lattice vector $\mathbf{g}'$ and $\mathbf{g}''$ label the cells where the i-th and j-th Aos are centered. Because of translation invariance of the integrals in the local basis, χ_i can be referred to the $\mathbf{0}$-cell, so that

$H_{ij}(\mathbf{g}) = \langle \chi_i(\mathbf{r} - \mathbf{r}_i) | H | \chi_j(\mathbf{r} - \mathbf{r}_j - \mathbf{g}) \rangle$

where ***g=g"-g'*** is a generic direct lattice vector.
The generic element of the Fock/Kohn-Sham matrix represented in the reciprocal space is:

$$H_{ij}(\mathbf{k}) = \langle\Phi_i(\mathbf{k};\mathbf{r})|H|\Phi_j(\mathbf{k};\mathbf{r})\rangle = 1/N \sum_{g'} \sum_{g} e^{i\mathbf{k}\cdot(\mathbf{g}-\mathbf{g}')} \langle\chi_i(\mathbf{r}-\mathbf{r}_i-\mathbf{g}')|H|\chi_j(\mathbf{r}-\mathbf{r}_j-\mathbf{g}'')\rangle$$

$$H_{ij}(\mathbf{k}) =' \sum_{g} e^{i\mathbf{k}\cdot\mathbf{g}} \langle\chi_i(\mathbf{r}-\mathbf{r}_i)|H|\chi_j(\mathbf{r}-\mathbf{r}_j-\mathbf{g})\rangle = \sum_{g} e^{i\mathbf{k}\cdot\mathbf{g}} H_{ij}(\mathbf{g})$$

The use of a localized basis set allows one:
(i) to efficiently evaluate the Fock/Kohn-Sham matrix in the direct space by means of standard molecular techniques and to fully exploit the point and translation symmetry of the crystal;
(ii) in a periodic approach, 2D (slabs), 1D (polymers) and 0D (molecules) model systems can be consistently treated by exploiting the translational symmetry along two-, one- and zero-directions, thus avoiding artificial 3D periodic reproduction of the model system,
(iii) a "chemical" interpretation of bonds as commonly carried out for molecules.
As stated before, results for 3D periodic structures are completely general and can be extended to lower dimensionality models by simply removing translational symmetry along one or two directions.
For further details and examples the reader is advised to refer to the review by Dovesi et al. [2].

Acknowledgements
I would like to acknowledge several people who introduced me to the world of solid state chemistry (C. Pisani, R. Dovesi, R. Orlando and C. Roetti) and crystallography (D. Viterbo and P. Ugliengo) for many fruitful discussions, sharing notes and slides.

References

[1] C. Pisani (Ed.) *Quantum Mechanical Calculation of the Properties of Crystalline Materials*, Springer-Verlag, Berlin, 1996
[2] R. Dovesi, B. Civalleri, R. Orlando, C. Roetti, V. R. Saunders, *Ab Initio Quantum Simulation in Solid State Chemistry*, Rev. Comp. Chem. 21 (2005) 1-125
(https://media.wiley.com/product_data/excerpt/9X/04716823/047168239X.pdf)
[3] C. Pisani "*Loss of Symmetry in Crystals: Surfaces and Local Defects*" in ref. [1]
[4] CRYSTAL Tutorials: *How to model surfaces: the slab model* -
http://tutorials.crystalsolutions.eu/tutorial.html?td=surfaces&tf=surfaces_tut
[5] CRYSTAL Tutorials: M. D'Amore, M. Causà: Nanotube systems -
http://tutorials.crystalsolutions.eu/tutorial.html?td=tuto_nano&tf=nano_md_g2
[6] CRYSTAL Tutorials: Y. Noel, R. Demichelis: Nanotube systems -
http://tutorials.crystalsolutions.eu/tutorial.html?td=nanotube&tf=nanotube
[7] D. Viterbo "*Crystal Lattices and Crystal symmetry*" in ref [1]
[8] R. Dovesi "*The language of Band Theory*" in ref. [1]
[9] R. Hoffman, *Solids and Surfaces: A Chemist's View of Bonding in Extended Structures*, VCH publisher, 1988

From Energy and Wavefunction to Advanced Properties of Solids

Alessandro Erba
Chemistry Department, University of Turin, Turin, Italy
alessandro.erba@unito.it

Quantum-mechanical simulations aim at solving the Schrödinger equation of the system. In most cases, the static (time-independent) equation is considered for the ground-state of the system:

$$\mathcal{H}\Psi = E\Psi$$

where $\mathcal{H}$ is the Hamiltonian operator (often approximated with the so-called "electrostatic, non-relativistic" Hamiltonian). The many-body electron-electron repulsion term of the Hamiltonian requires special care and is often described by means of a mean-field Hartree-Fock (HF) scheme (where each electron feels the average field of all the others) or a density functional theory (DFT) approach or a combination of the two (hybrid DFT). Another common approximation in condensed matter simulations consists in using a single-determinant to represent the ground state wave-function Ψ.

The Born-Oppenheimer approximation (BOA) is also adopted according to which, due to their different velocity, the motion of electrons can be decoupled from the motion of nuclei. At each fixed nuclear configuration R, the electronic Schrödinger equation is solved:

$$\mathcal{H}(R)\Psi(R) = E(R)\Psi(R) \tag{1}$$

which provides the electronic energy $E(R)$ and the ground state wave-function $\Psi(R)$ of the considered nuclear configuration. The BOA basically assumes that electrons would instantaneously rearrange to a change in nuclear configuration.

In these notes, we do not discuss how the equation (1) is solved in a solid state context (in this respect, see lecture notes by B. Civalleri and P. Giannozzi). Here, we just assume that we know how to solve that equation for each nuclear configuration R and we provide an overview (far from being complete) of what properties of materials can be computed from there. First of all, we may define two groups of properties: those obtained from the analysis of the energy $E(R)$ and those obtained from the analysis of the wave-function $\Psi(R)$. In particular, we will give more details about the first group because the properties belonging to the second group will also be covered by many other lecture notes.

Energy-related Properties

Many properties can be extracted from the analysis of the energy itself while many others from the study of its derivatives with respect to different kinds of perturbations. Here, we consider three types of perturbations: atomic displacements u, lattice parameters a and electric field components $\mathcal{E}$. A systematic formal account on the properties that can be obtained from energy derivatives with respect to these perturbations can be found in the references below up to second- and third-order, respectively:

1) X. Wu, D. Vanderbilt and D.R. Hamann, *Phys. Rev. B*, **72**, 035105 (2005)

2) M. Veithen, X. Gonze and P. Ghosez, *Phys. Rev. B*, **71**, 125107 (2005)

The Potential Energy (Hyper)Surface

However, before discussing its derivatives, let us analyze the energy $E(R)$ as such. In order to do that, it proves convenient to introduce the fundamental concept of a potential energy (hyper)surface (PES), which basically describes the energy of the system as a function of its nuclear configuration:

$$E = f(v_1, v_2, \cdots, v_M)$$

The PES is a (hyper)surface in a M-dimensional space where $v_1, v_2, \cdots, v_M$ are the variables used to define the structure of the system (atomic positions and lattice parameters for a solid). Different sets of coordinates can be used (Cartesian, internal, etc.) so that there is no upper bound for M (redundant coordinates can be used). However there is a lower bound for M (i.e. minimum number of variables needed to define the structure). For a molecule it is:

$$M = 3N - 3 - 3(2)$$

where N is the number of atoms and where the translational and rotational degrees of freedom are removed. For a 3D solid it is:

$$M = 3N - 3 + (9 - 3)$$

where now N is the number of atoms per cell and the $+9$ refers to the number of variables needed to define the lattice cell.

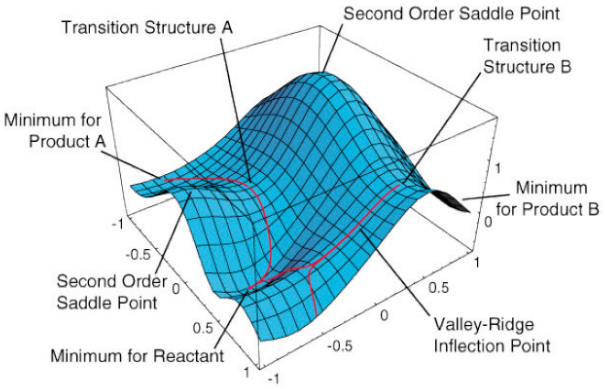


Figure 1: Schematic representation of the potential energy (hyper)surface and of its critical points.

Some critical points of the PES embody relevant chemical information. In particular, minima of the PES correspond to equilibrium chemical structures while first-order saddle points to transition state structures. All critical points are characterized by a vanishing gradient vector $\mathbf{g}$:

$$\mathbf{g} = \left(\frac{\partial E}{\partial v_1}, \frac{\partial E}{\partial v_2}, \cdots, \frac{\partial E}{\partial v_M} \right)$$

Critical points can be minima, maxima or saddle points depending on the sign of the eigenvalues of the Hessian matrix of second energy derivatives:

$$\mathbf{H} \rightarrow H_{ij} = \frac{\partial^2 E}{\partial v_i \partial v_j} \qquad \mathbf{HU} = \mathbf{\Lambda U}$$

Once diagonalized, the Hessian matrix provides M eigenvalues λ_i: if they are all positive then the critical point is a minimum; if they are all negative it is a maximum; if they are all positive but n negative then it is a n-th order saddle point.

Geometry Optimization

The process of finding minima (or saddle points) of the PES (i.e. geometry optimization) is crucial for the study of many advanced properties of a system. Several schemes can be devised of increasing complexity, cost and efficiency depending on the "ingredients" used in the minimization process (energy, forces, Hessian). In particular, the evaluation of the Hessian is much more expensive than that of the forces, which in turn is more expensive than that of the energy:

$$E < \mathbf{g} \ll \mathbf{H}$$

Energy-based methods (line optimization, etc.) are simple, cheap, widely applicable but typically show very slow convergence. Gradient-based methods (conjugate gradient, quasi-Newton, etc.) are more expensive but show good convergence. Second derivative methods (Newton, Newton-Raphson, etc.) are very expensive but also very fast. Most solid state packages implement gradient-based algorithms for geometry optimization based on the analytical evaluation of the forces.

Let us discuss into some detail the Newton (and quasi-Newton) scheme. In this case, the PES is approximated as a quadratic function (Taylor expansion centered around an initial configuration 0 and truncated to second-order):

$$E(\mathbf{v}) = E_0 + \mathbf{g}_0^\top \Delta\mathbf{v} + \frac{1}{2}\Delta\mathbf{v}^\top \mathbf{H} \Delta\mathbf{v}$$

where $\Delta\mathbf{v} = \mathbf{v} - \mathbf{v}_0$ is the difference between current and starting configurations. The gradient vector at the current configuration **v** can thus be written as:

$$\mathbf{g}(\mathbf{v}) = \mathbf{g}_0 + \mathbf{H}\Delta\mathbf{v}$$

Given that we look for a minimum, we can impose the following condition:

$$\mathbf{g}(\mathbf{v}) = \mathbf{g}_0 + \mathbf{H}\Delta\mathbf{v} = \mathbf{0}$$

which leads to:

$$\Delta\mathbf{v} = -\mathbf{H}^{-1}\mathbf{g}_0$$

This is known as Newton's step and tells us how the geometry needs to be changed to pass from an initial configuration to a minimum, provided that the PES is quadratic and that both the forces and Hessian of the initial configuration are known. In this case, a single optimization step would take us to the minimum. However, the PES is never exactly quadratic; furthermore, the numerical evaluation of the Hessian would be very expensive so that in quasi-Newton approaches the Hessian is never explicitly computed but rather estimated from the evolution of the forces along the optimization process (for instance with the Broyden-Fletcher-Goldfarb-Shanno, BFGS, approach).

Pressure

The effect of pressure can be included on computed thermodynamic, structural and mechanical properties by performing constrained geometry optimizations:

- Pressure-constrained optimizations - An external pressure is applied in the form of a hydrostatic prestress and the lattice parameters (and atomic positions) optimized under this constraint. In this case, the quantity being minimized during the optimization process is the enthalpy $H = E + PV$ of the system. The optimization provides the equilibrium volume (and structure) at the desired pressure;

- Volume-constrained optimizations - Several constant-volume optimizations are performed at different volumes so that an energy-volume curve is obtained and fitted to an equation-of-state. The pressure-volume relation $P(V)$ and the value of the bulk modulus K of the system (i.e. average elastic response of the system) are then obtained from the following equations:

$$P(V) = -\left(\frac{\partial E}{\partial V}\right) \quad \text{and} \quad K(V) = V\left(\frac{\partial^2 E}{\partial V^2}\right)$$

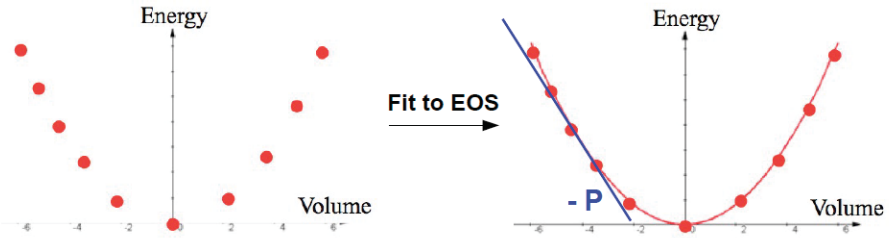


Figure 2: Schematic representation of the equation-of-state approach to the determination of the $P(V)$ relation.

Harmonic Frequencies

Atoms in a lattice are never at rest in their equilibrium positions but they rather vibrate because of thermal nuclear motion (also at 0 K they vibrate because of zero-point motion). The lattice dynamics can be described at the harmonic level (i.e. with a quadratic description of the PES with respect to atomic displacements). In this case, the nuclear motion is described in terms of $3N-3$ independent quantum harmonic oscillators. A detailed description of lattice vibrations is crucial to the understanding of infrared (IR) and Raman spectroscopies, thermodynamics, thermal properties, transport properties, phase-transitions, nuclear-relaxation, etc. Harmonic vibration frequencies ω_i are obtained from the diagonalization of the mass-weighted Hessian matrix of second energy derivatives with respect to pairs of atomic displacements ($\omega_i = \sqrt{\lambda_i}$):

$$\mathbf{W} \rightarrow W_{ij} = \frac{H_{ij}}{\sqrt{M_i M_j}} \quad \text{where} \quad H_{ij} = \frac{\partial^2 E}{\partial u_i \partial u_j} \quad \text{then} \quad \mathbf{WU} = \mathbf{\Lambda U}$$

where i and j represent Cartesian atomic displacements. The $\mathbf{U}$ eigenvector matrix contains the normal modes of vibration. The Hessian matrix $\mathbf{H}$ is often computed numerically from finite differences of analytical gradients in solid state packages.

If only relative motions of atoms within the same cell are considered and translational invariance retained, then in-phase (Γ-point) atomic motions are described that can be probed by IR and Raman spectroscopies:

$$\mathbf{W}^{\Gamma} \rightarrow W^{\Gamma}_{ij} = \frac{H^{\mathbf{0}}_{ij}}{\sqrt{M_i M_j}} \qquad H^{\mathbf{0}}_{ij} = \frac{\partial^2 E}{\partial u_{i\mathbf{0}} \partial u_{j\mathbf{0}}}$$

where bold symbols label lattice vectors. Harmonic frequencies provide the positions of IR and Raman peaks. Also the intensity of the peaks can be computed from energy derivatives. In particular, the IR intensity (i.e. variation of the dipole moment along the normal mode) can be obtained from the second-rank Born charge tensor $\mathbf{Z}^*$ of mixed second energy derivatives:

$$\mathcal{I}^{\text{IR}} = \mathcal{I}^{\text{IR}}(\mathbf{Z}^*) \quad \text{where} \quad Z^*_{ij} = \frac{\partial^2 E}{\partial \mathcal{E}_i \partial u_j}$$

The Born tensors can be computed in many different ways: in terms of localized crystalline orbitals (Wannier functions), from the Berry phase, from coupled-perturbed-Hartree-Fock/Kohn-Sham (CPHF/KS) calculations, from the density functional perturbation theory (DFPT), etc. The Raman intensity (i.e. variation of the polarizability along the normal mode) is obtained from the third-rank Raman tensors $\mathbf{A}$:

$$\mathcal{I}^{\text{Raman}} = \mathcal{I}^{\text{Raman}}(\mathbf{A}) \quad \text{where} \quad A_{ijk} = \frac{\partial^3 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j \partial u_k}$$

If phase modulations of atomic motions in different lattice cells are considered, then phonon dispersion is taken into account:

$$\mathbf{W}^{\mathbf{k}} \rightarrow W_{ij}^{\mathbf{k}} = \sum_{\mathbf{g}} \frac{H_{ij}^{\mathbf{g}}}{\sqrt{M_i M_j}} e^{i\mathbf{k}\cdot\mathbf{g}} \qquad H_{ij}^{\mathbf{g}} = \frac{\partial^2 E}{\partial u_{i\mathbf{0}} \partial u_{j\mathbf{g}}}$$

By diagonalization of the dynamical matrices $\mathbf{W}^{\mathbf{k}}$ one gets the phonon frequencies at each $\mathbf{k}$ point. A full account of phonon dispersion is crucial to the description of many thermodynamic properties, to the calculation of accurate phonon density-of-states, inelastic neutron scattering spectra, etc.

Tensorial Properties of Crystals

Many properties of crystals require a tensorial representation because of their anisotropic structure. For instance, while a scalar number (the dielectric constant) is enough to describe the propagation velocity of an electric field in a liquid, the 9 elements of a second-rank 3×3 tensor are needed for a crystal (i.e. light propagates with different velocity along different crystallographic directions).

Linear and Non-linear Optical Properties

These properties are due to the interaction between light and matter. Light is an electromagnetic radiation with an oscillating electric field and an orthogonal oscillating magnetic field. Several approximations can be introduced to treat such a complex phenomenon. First of all, one might consider the electric field only:

$$\boldsymbol{\mathcal{E}}(\mathbf{r}, \omega) = \boldsymbol{\mathcal{E}}_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega\mathbf{t})}$$

The field can then be considered as static (time-independent, $\omega = 0$):

$$\boldsymbol{\mathcal{E}}(\mathbf{r}) = \boldsymbol{\mathcal{E}}_0 e^{i(\mathbf{k}\cdot\mathbf{r})}$$

If UV-visible radiation is considered then the wavelength of the electric field is much longer then typical inter-atomic distances so that the field can also be considered as constant ($\mathbf{k} = \mathbf{0}$):

$$\boldsymbol{\mathcal{E}}(\mathbf{r}) = \boldsymbol{\mathcal{E}}_0$$

From an expansion of the energy in a Taylor's series with respect to the electric field

$$E(\boldsymbol{\mathcal{E}}) = E(\mathbf{0}) + \boldsymbol{\mu}\cdot\boldsymbol{\mathcal{E}} + \frac{1}{2}\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{E}}^2 + \frac{1}{3!}\boldsymbol{\beta}\cdot\boldsymbol{\mathcal{E}}^3 + \frac{1}{4!}\boldsymbol{\gamma}\cdot\boldsymbol{\mathcal{E}}^4$$

we introduce several static optical properties as energy derivatives (of increasing order) with respect to electric field components:

$$\begin{aligned}
\text{dipole moment} \quad \mu_i &= \left(\frac{\partial E}{\partial \mathcal{E}_i}\right) \\
\text{polarizability} \quad \alpha_{ij} &= \left(\frac{\partial^2 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j}\right) \\
\text{first hyper-polarizability} \quad \beta_{ijk} &= \left(\frac{\partial^3 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j \partial \mathcal{E}_k}\right) \\
\text{second hyper-polarizability} \quad \gamma_{ijkl} &= \left(\frac{\partial^4 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j \partial \mathcal{E}_k \partial \mathcal{E}_l}\right)
\end{aligned}$$

From the polarizability, one gets the dielectric tensor as:

$$\boldsymbol{\epsilon} = 1 + \frac{4\pi}{V}\boldsymbol{\alpha}$$

where V is the cell volume. We remind that the refractive index is proportional to the square root of the dielectric response $n \propto \sqrt{\epsilon}$. Linear and non-linear electric susceptibilities can be easily obtained from these computed quantities.

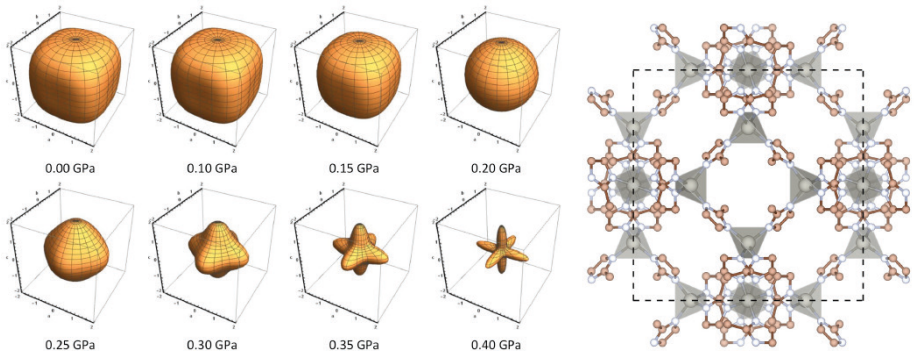


Figure 3: 3D representation of the pressure dependence of the directional Young modulus of ZIF-8.

Mechanical/Elastic Properties

The stress-strain relation for a spring is given by Hook's law $f = -kx$, where f is the applied force, x the induced elongation and k is the spring elastic constant. Because of the anisotropic structure of a crystal, both stress and strain become second-rank symmetric tensors so that the generalization of Hook's law to the case of an anisotropic medium reads:

$$\sigma_{ij} = \sum_k \sum_l C_{ijkl}\eta_{kl}$$

where $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{\eta}$ is the strain tensor, and the C_{ijkl} (elastic constants) are the elements of a fourth-rank elastic tensor providing the link between stress and strain. It can be shown that the elastic constants can be defined as second energy derivatives with respect to pairs of lattice deformations:

$$C_{ijkl} = \frac{1}{V}\left(\frac{\partial^2 E}{\partial \eta_{ij} \partial \eta_{kl}}\right)$$

These quantities are computed from either i) numerical second energy derivatives, or ii) numerical finite differences of analytical cell gradients. Many elastic/mechanical properties can then be easily obtained from the computed elastic tensor: bulk modulus, Young modulus, shear modulus, Poisson's ratio, elastic wave velocities, etc.

Piezoelectricity

Piezoelectricity is the ability of certain materials to convert a mechanical force into an electrical signal, or viceversa. The *direct* piezoelectric effect (described by a third-rank tensor **e**) consists in applying a mechanical deformation to a non-centrosymmetric crystal and in inducing an electric polarization. The *converse* piezoelectric effect (described by a third-rank tensor **d**) consists in applying an external electric field to a non-centrosymmetric crystal and in inducing a structural deformation. The direct piezoelectric constants can be defined as derivatives of the polarization with respect to a strain or, equivalently, as second mixed energy derivatives with respect to electric field components and strain components:

$$e_{kij} = \left(\frac{\partial P_k}{\partial \eta_{ij}}\right) = \frac{1}{V}\left(\frac{\partial^2 E}{\partial \mathcal{E}_k \partial \eta_{ij}}\right)$$

They can be computed from the Berry phase, from CPHF/KS calculations, from DFPT, etc. Once the direct piezoelectric constants and the elastic constants are computed, one can obtain the converse piezoelectric ones from their combination as follows:

$$\mathbf{d} = \mathbf{eS}$$

where $\mathbf{S} = \mathbf{C}^{-1}$ is the elastic compliance tensor.

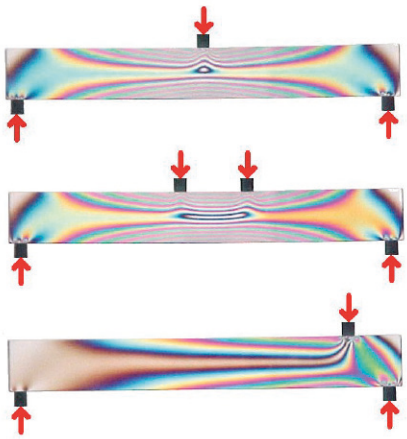


Figure 4: Stress-optical effect.

Photoelasticity and Piezo-optics

The variation of the elements of the dielectric tensor (linked to the refractive indices) with respect to internal or applied strain constitutes the so-called photoelasticity, or elasto-optics, of a crystal. The photo-elastic Pockels constants are the elements of a fourth-rank tensor linking the variation of the inverse of the dielectric

tensor to the elements of the strain tensor and can be computed as third mixed energy derivatives:

$$p_{ijkl} = \left(\frac{\partial \epsilon_{ij}}{\partial \eta_{kl}}\right) = \frac{4\pi}{V}\left(\frac{\partial^3 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j \partial \eta_{kl}}\right)$$

From the knowledge of the photo-elastic tensor **p** and of the elastic compliance tensor **S**, the piezo-optic tensor can also be computed,

$$\boldsymbol{\pi} = \mathbf{pS} \quad \text{with} \quad \mathbf{S} = \mathbf{C}^{-1}$$

which provides useful information on the dependence on stress of the linear optical properties of a crystal.

Wavefunction-related Properties

Many properties of a system, related to the electron distribution in position and momentum spaces, can be computed from the analysis of its wavefunction. However, the wavefunction is a rather complex quantity to deal with. For a system of N electrons, it is a function of N space-spin variables:

$$\Psi(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N) \quad \text{with} \quad \mathbf{X} \equiv (\mathbf{r}, \omega)$$

where **r** are the electron spatial coordinates and ω is the electron spin variable. It proves convenient to introduce so-called density matrices to simplify the analysis of the wavefunction. The m-order density matrix is a function of $2m$ space-spin variables defined as:

$$\Gamma^m(\mathbf{X}_1, \ldots, \mathbf{X}_m; \mathbf{X}'_1, \ldots, \mathbf{X}'_m) =$$

$$\int \Psi(\mathbf{X}_1, \ldots, \mathbf{X}_m, \ldots, \mathbf{X}_N)\Psi^*(\mathbf{X}'_1, \ldots, \mathbf{X}'_m, \ldots, \mathbf{X}_N) d\mathbf{X}_{m+1} \ldots d\mathbf{X}_N$$

Density matrices play a very important role in the analysis of the electron distribution because of the following theorem:

"The expectation value of any m-particle operator can be obtained from the m-order density matrix $\boldsymbol{\Gamma}^m$"

In particular, one-electron properties (such as the electron charge distribution, the electron momentum distribution, X-ray structure factors, Compton profiles, auto-correlation functions, the electrostatic potential, the kinetic energy, etc.) can thus be computed from the first-order density matrix $\Gamma^1(\mathbf{X}; \mathbf{X}')$ that is a function of just 2 space-spin variables. A schematic representation is given in Figure 5 of the connections among the first-order density matrix and some one-electron properties in position and momentum spaces.

Here we are not going to discuss into detail how these properties can be computed and analyzed because several other lectures cover these topics. We just provide a partial list of the properties that can be obtained from the wavefunction of the system, and of the schemes for their analysis.

Electron Charge Density Analysis

The electron charge distribution embodies rich information on the chemical nature of the interactions taking place in the system. A variety of techniques can be used to try to extract as much information as possible out of it.

A visual analysis can be performed based on 2D maps or 3D plots. The total electron density can be represented as well as different "deformation densities" defined as the difference between the total density

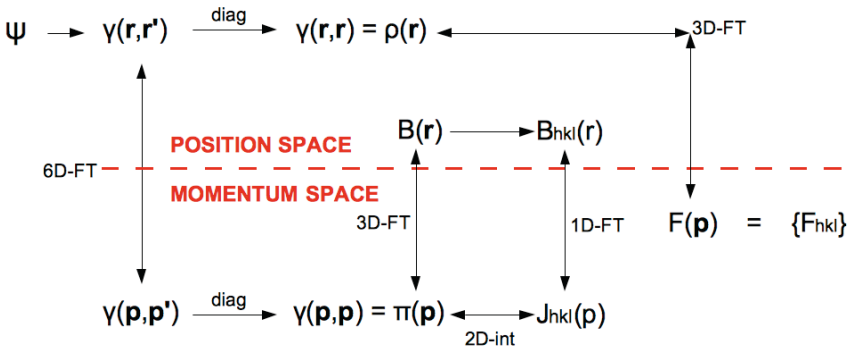


Figure 5: Schematic representation of the connections among the first-order density matrix and some one-electron properties in position and momentum spaces.

and some reference densities (i.e. atomic densities of non-interacting atoms or ions, densities of non-interacting molecules, etc.). Colored 3D representations of the density can be obtained by taking into account the local value of the electrostatic potential, which allows to identify possible chemically active and accessible sites to guest molecules in complex structures. For open-shell, magnetic systems, the spin density $\sigma(\mathbf{r}) = \rho^{\alpha}(\mathbf{r}) - \rho^{\beta}(\mathbf{r})$ can also be computed and plotted.

A large variety of schemes can be used to analyze the electron charge density numerically. X-ray structure factors can be computed and compared to the experiment (dynamical thermal effects can be included by computing Debye-Waller factors from atomic anisotropic displacement parameters). The total electron charge density can be partitioned into atomic contributions according to several schemes (Mulliken, Born, Bader, Hirshfeld, etc.). A topological analysis can be performed as in Bader's quantum theory of atoms in molecules (QTAIM). The spatial distribution of the density can be analyzed in terms of the Electron Localization Function (ELF), Maximum Probability Domain (MPD), and many other schemes.

Electron Momentum Density Analysis

The electron momentum density provides a complementary information with respect to the electron charge density. As per the electron charge density, 2D maps or 3D plots can be used to visually analyze it. At variance with the electron charge density, the electron momentum density is a function centered in a single point (the origin of momentum space) so that the study of its anisotropy is crucial to its analysis. The total density can be plotted as well as the difference with respect to its "spherical average".

Directional Compton profiles and auto-correlation functions are properties related to the electron momentum distribution that can be computed. For open-shell systems, magnetic Compton profiles can also be evaluated as differences between contributions from spin-up and spin-down electrons.

Introduction to density-functional theory and the plane-wave pseudopotential method

Paolo Giannozzi
Department of Mathematics, Computer Science, and Physics, University of Udine, Udine, Italy
paolo.giannozzi@uniud.it

Density Functional Theory (DFT) is a ground-state theory in which the emphasis is on the charge density as the relevant physical quantity. DFT is highly successful in describing structural and electronic properties in a vast class of materials, ranging from atoms and molecules to simple crystals to complex extended systems (including glasses and liquids). DFT has become a common tool in first-principles calculations aimed at describing - or even *predicting* - properties of molecular and condensed matter systems.

The Hohenberg-Kohn theorem Let us consider a system of N interacting spinless electrons under an external potential $V(\boldsymbol{r})$ (usually the Coulomb potential of the nuclei). If the system has a nondegenerate ground state, it is obvious that there is only one ground-state charge density $n(\boldsymbol{r})$ that corresponds to a given $V(\boldsymbol{r})$. In 1964 Hohenberg and Kohn demonstrated the opposite, far less obvious result: there is only one external potential $V(\boldsymbol{r})$ that yields a given ground-state charge density $n(\boldsymbol{r})$. The demonstration is very simple, using a *reductio ad absurdum* argument. Let us consider a many-electron Hamiltonian $H=T+U+V$, with ground state wavefunction Ψ. T is the kinetic energy, U the electron-electron interaction, V the external potential. The charge density $n(\boldsymbol{r})$ is defined as $n(\boldsymbol{r})= N \int|\Psi(\boldsymbol{r},\boldsymbol{r}_2, \boldsymbol{r}_3,...,\boldsymbol{r}_N)|^2 d\boldsymbol{r}_2...d\boldsymbol{r}_N$. Let us consider now a *different* Hamiltonian $H'=T+U+V'$ (V and V' do not differ simply by a constant), with ground state wavefunction Ψ'. Let us assume that the ground state charge densities are the same: $n[V]=n'[V']$. The following inequality holds: $E'=\langle\Psi'|H'|\Psi'\rangle < \langle\Psi|H'|\Psi\rangle$, but $\langle\Psi|H'|\Psi\rangle=\langle\Psi|H-V+V'|\Psi\rangle$, that is, $E' < E + \int (V'(\boldsymbol{r})-V(\boldsymbol{r}))n(\boldsymbol{r})d\boldsymbol{r}$. The inequality is strict because Ψ and Ψ' are different eigenstates of different Hamiltonians. By reversing the primed and unprimed quantities, one obtains an absurd result. This demonstrates that no two different potentials can have the same charge density.

The Kohn-Sham equations One year later, Kohn and Sham (KS) reformulated the problem in a more familiar form and opened the way to practical applications of DFT. The system of interacting electrons is mapped on to an auxiliary system of non-interacting electrons having the same ground state charge density $n(\boldsymbol{r})$. For a system of non-interacting electrons the ground-state charge density is representable as a sum over one-electron orbitals, the KS orbitals $\psi_i(\boldsymbol{r})$: $n(\boldsymbol{r})=2\Sigma_i |\psi_i(\boldsymbol{r})|^2$, where i runs from 1 to $N/2$ if we assume double occupancy of all states, and the KS orbitals are the solutions of the Schrödinger equation

$$(- (\hbar^2/2m)\nabla^2 + V_{KS}(\boldsymbol{r}))\psi_i(\boldsymbol{r})=\varepsilon_i \psi_i(\boldsymbol{r})$$

(m is the electron mass) obeying orthonormality constraints $\int \psi^*_i(\boldsymbol{r})\psi_j(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = \delta_{ij}$. The existence of a unique potential $V_{KS}(\boldsymbol{r})$ having $n(\boldsymbol{r})$ as its ground state charge density is a consequence of the Hohenberg-Kohn theorem, holding irrespective of the electron-electron interaction U.
The problem to determine $V_{KS}(\boldsymbol{r})$ for any given $n(\boldsymbol{r})$ is solved variationally: for an arbitrary variation of the $\psi_i(\boldsymbol{r})$, under the orthonormality constraints, the first-order variation of E must vanish. This implies that the functional derivative, $\delta E'/\delta\psi_i(\boldsymbol{r})$, of the constrained functional, $E' = E - \sum_{ij} \lambda_{ij} (\int \psi^*_i(\boldsymbol{r})\psi_j(\boldsymbol{r})\mathrm{d}\boldsymbol{r} - \delta_{ij})$, where λ_{ij} are Lagrange multipliers, must vanish. It is convenient to rewrite the energy functional as follows:

$$E = T_s[n(\boldsymbol{r})] + \mathrm{E}_H[n(\boldsymbol{r})] + \mathrm{E}_{XC}[n(\boldsymbol{r})] + \int n(\boldsymbol{r})V(\boldsymbol{r})\mathrm{d}\boldsymbol{r}\,.$$

The first term is the kinetic energy of *non-interacting* electrons:

$$T_s[n(\boldsymbol{r})] = -(\hbar^2/2m)\, 2\sum_i \int \psi^*_i(\boldsymbol{r})\nabla^2\psi_j(\boldsymbol{r})\mathrm{d}\boldsymbol{r}\,.$$

The second term (called the Hartree energy) contains the electrostatic interactions between clouds of charge:

$$\mathrm{E}_H[n(\boldsymbol{r})] = (\mathrm{e}^2/2)\int n(\boldsymbol{r})|\boldsymbol{r}-\boldsymbol{r}'|^{-1}n(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'.$$

The third term, called *exchange-correlation* (XC) *energy*, contains all the remaining terms and our ignorance. The logic behind such procedure is to subtract out easily computable terms which account for a large fraction of the total energy. Using $\delta n(\boldsymbol{r})/\delta\psi^*_i(\boldsymbol{r}) = \psi^*_i(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r})$ one finds: $\delta T_s/\delta\psi^*_i(\boldsymbol{r}) = -(\hbar^2/2m)2\sum_i \nabla^2\psi_i(\boldsymbol{r})$; the *Hartree* potential $V_H(\boldsymbol{r}) = \delta E_H/\delta\psi^*_i(\boldsymbol{r}) = \mathrm{e}^2\int n(\boldsymbol{r}')|\boldsymbol{r}-\boldsymbol{r}'|^{-1}\mathrm{d}\boldsymbol{r}'\psi_i(\boldsymbol{r})$, the XC potential $V_{XC}(\boldsymbol{r}) = (\delta E_{XC}/\delta n(\boldsymbol{r}))$ By making a subspace rotation in the space of ψ_i's that leaves the charge density invariant, we can bring the matrix of Lagrange multipliers to diagonal form: $\lambda_{ij} = \delta_{ij}\varepsilon_i$. We finally obtain the KS equations:

$$\left(-(\hbar^2/2m)\nabla^2 + V_H(\boldsymbol{r}) + V_{XC}[n(\boldsymbol{r})] + V(\boldsymbol{r})\right)\psi_i(\boldsymbol{r}) = \varepsilon_i\,\psi_i(\boldsymbol{r}),$$

or $(H_{KS}-\varepsilon_i)\psi_i(\boldsymbol{r}) = 0$, where $H_{KS} = (-\hbar^2\nabla^2/2m) + V_{KS}(\boldsymbol{r})$ is the KS Hamiltonian, $V_{KS} = V_H(\boldsymbol{r}) + V_{XC}(\boldsymbol{r}) + V(\boldsymbol{r})$. Note that $H_{KS}(\boldsymbol{r})$ is related to the functional derivative of the energy via $\delta E/\delta\psi^*_i(\boldsymbol{r}) = H_{KS}\psi_i(\boldsymbol{r})$.

Exchange-correlation functional To be useful, KS equations require an approximation to the (unknown) E_{XC} and V_{XC} functionals. As early as 1965 KS introduced the *Local Density Approximation* (LDA): in the spirit of Slater's local exchange, E_{XC} is approximated using the XC energy density $\varepsilon_{XC}(n)$ of the homogeneous electron gas at the local density $n = n(\boldsymbol{r})$:

$$E_{XC}[n(\boldsymbol{r})] = \int n(\boldsymbol{r})\,\varepsilon_{XC}(n(\boldsymbol{r}))\,\mathrm{d}\boldsymbol{r}$$

Approximate forms for $\varepsilon_{XC}(n)$ have been known for a long time, going back to Wigner (1931). Highly accurate results from Quantum Monte-Carlo techniques were found by Ceperley and Alder and parameterized by Perdew and Zunger:

$$\varepsilon_{XC}(n) = -0.9164/r_s - 0.2846/(1+1.0529\, r_s^{1/2} + 0.3334\, r_s),\ r_s \geq 1;$$

$$\varepsilon_{XC}(n) = -0.9164/r_s - 0.0960 + 0.0622\log(r_s) - 0.0232\, r_s + 0.0040\, r_s\log(r_s),\ r_s \leq 1.$$

Here $r_s = (3n/4\pi)^{1/3}$ is in Bohr radii and ε_{XC} in Ry. The first term is the Hartree-Fock exchange contribution, the remaining terms are *correlation* energy. The XC potential is straightforwardly obtained by derivation: $V_{XC}(\boldsymbol{r}) = (\varepsilon_{XC}(n) + n(\mathrm{d}\varepsilon_{XC}(n)/\mathrm{d}n))_{n=n(\boldsymbol{r})}$.

In spite of its simplicity, LDA is surprisingly successful: bond lengths, lattice parameters, elastic constants and vibrational frequencies, surface energies for many condensed-matter systems are computed with good accuracy, down to 1÷2% errors. There are however also some serious drawbacks, notably:

- The (in)famous *band gap problem*: the gap computed as $\Delta_g=\varepsilon_c-\varepsilon_v$ (or HOMO-LUMO in quantum chemistry parlance) wildly underestimates (~50%) the true band gap
- Inability to properly describe *strongly correlated* materials, e.g., transition metal oxides
- Strongly overestimated (~20%) cohesive energies, sometimes hiding a fundamental problem: van der Waals (dispersive) interactions are in principle absent from LDA.

Our understanding of the origin of such problems has much progressed in recent years. It can be shown that the exact functional must obey *sum rules* and have some features that are hard to reproduce with LDA. In particular: i) If we consider the extension to fractional charge, V_{XC} must have a *discontinuity* for integer values; ii) the *self-interaction* of an electron with itself must be absent (it is in Hartree-Fock, not so in LDA).

The search for better functionals is an active research field, producing a flurry of improved functionals that yield better and better results, at the price of an increased computational complexity. John Perdew introduced the *Jacob's Ladder* metaphor:

1. At the lowest rung, LDA, with $\varepsilon_{XC}(n(\boldsymbol{r}))$ depending only upon the local charge density.
2. The next rung is the Generalized Gradient Approximation (GGA), in which ε_{XC} depends also upon the gradient of $n(\boldsymbol{r})$: $E_{XC}=\int n(\boldsymbol{r})\varepsilon_{XC}(n(\boldsymbol{r}),|\nabla n(\boldsymbol{r})|)\mathrm{d}\boldsymbol{r}$. With respect to LDA, GGA is marginally more expensive, but yields better structural properties and much better cohesive energies. It doesn't solve any of the other major problems cited above, though. GGA is the basic "standard" in DFT simulations since ~20 years and until recently. Some of the most popular GGA's are PBE, PW91, BLYP.
3. One rung up, the recent "meta-GGA" functionals: $E_{XC}=\int n(\boldsymbol{r})\varepsilon_{XC}(n(\boldsymbol{r}),|\nabla n(\boldsymbol{r})|,\tau(\boldsymbol{r}))\mathrm{d}\boldsymbol{r}$, where $\tau(\boldsymbol{r}) = \sum_i |\nabla \psi_i(\boldsymbol{r})|^2$ is the "kinetic energy density". These functionals are more complex to use: the KS Hamiltonian does not have any longer a simple form, more computationally expensive and numerically difficult than plain LDA or GGA. They are however able to deal with different types of bonding and yield very accurate results for structural properties. The recent SCAN functional is a promising candidate to replace GGA as "standard", but needs faster and more stable implementations than today's.
4. One rung up, *hybrid* functionals, containing some admixture of the exact exchange energy E_{HF}, calculated with KS orbitals as in Hartree-Fock theory:

   $$E_{XC}=\alpha_x E_{HF}+(1-\alpha_x)E_X+E_C$$

   where E_X and E_C are the exchange and correlation part of the DFT functional, respectively, and α_x=0.2÷0.3, depending upon the specific functional (PBE0: α_x=0.25, B3LYP: α_x=0.2). The XC potential contains a nonlocal V_x contribution. Hybrid functionals correct most of the GGA problems and yield very accurate results, that

however depend upon α_x, often used as an adjustable parameter. Hybrid functionals are easily and efficiently implemented using a localized basis set and are thus the method of choice for molecular calculations using quantum chemistry codes. With a plane-wave basis set, instead, the implementation is straightforward but very slow (tens of times more than GGA). Much progress is underway in this field with the usage of localized orbitals (Wannier functions and the like) to compute the V_x contribution.

5. One rung up: RPA-like approaches, very accurate and computational very heavy.

In addition to functionals mentioned above, it is worth mentioning a few "DFT+corrections" approaches for strongly correlated materials and for van der Waals forces:

- DFT+U: a Hubbard-like term, accounting for strong Coulomb correlations in systems with highly localized, atomic-like states, is added (typically to a GGA functional): $E_{\mathrm{DFT+U}}[n(\boldsymbol{r})] = E_{XC}[n(\boldsymbol{r})] + \mathrm{E}_U[n(\boldsymbol{r})]$, where $\mathrm{E}_U[n(\boldsymbol{r})] = U/2 \sum_\sigma \mathrm{Tr}[\boldsymbol{n}^\sigma(1-\boldsymbol{n}^\sigma)]$, U is a Coulomb repulsion (a few eV, system-dependent), $\boldsymbol{n}^\sigma$ is the matrix of orbital occupancies of the chosen set of atomic-like states. DFT+U is a quick-and-dirty but economical solution for a deep problem of DFT: the lack of discontinuity in approximated functionals, due to incomplete self-interactions cancellation, favors unphysical fractionary occupancy of localized states. DFT+U also improves the gap and level alignement in heterostructures, at the price of introducing an adjustable parameter U.
- vdW-DF functionals: E_{XC} contains a nonlocal term $E_{nl}[n(\boldsymbol{r})]=\frac{1}{2}\int n(\boldsymbol{r})\ \Phi(\boldsymbol{r},\boldsymbol{r}')\ n(\boldsymbol{r}')\mathrm{d}\boldsymbol{r}\ d\boldsymbol{r}'$, where $\Phi(\boldsymbol{r},\boldsymbol{r}') = \Phi(n(\boldsymbol{r}),\nabla n(\boldsymbol{r}),n(\boldsymbol{r}'),\nabla n(\boldsymbol{r}'),|\boldsymbol{r}-\boldsymbol{r}'|)$, that accounts for van der Waals forces. Can be computed with a reasonable computational overhead (Soler's technique) with respect to GGA, yielding generally good results, but sometimes overestimating binding.
- Grimme's DFT+D, adding a semi-empirical correction to GGA:

  $$E_{DFT+D} = E_{XC} + E_{vdW}(\boldsymbol{R}_\mathrm{i}), \qquad E_{vdW} = -(s_6/2) \sum_{i\neq j} (C_6^{\mathrm{ij}}/R^6_{\mathrm{ij}}) f_{damp}(R_{\mathrm{ij}})$$

  where the $\boldsymbol{R}_\mathrm{i}$ are atomic positions, $R_{\mathrm{ij}}=|\boldsymbol{R}_\mathrm{i}-\boldsymbol{R}_\mathrm{j}|$, s_6 is a global scaling factor depending upon the specific GGA, $C_6^{\mathrm{ij}}=(C_6^{\mathrm{i}}C_6^{\mathrm{j}})^{1/2}$, where C_6^{i} are dispersion coefficients for the i-th atom; the damping function f_{damp} prevents singularities for $R\rightarrow 0$. Parameters are fitted to experiments or to accurate theoretical data. The more sophisticated DFT+D3 version includes three-body and other terms in the semi-empirical correction, yielding good results for vdW-bonded system with a modest computational overhead with respect to GGA. It can hardly be considered a "first-principle" approach, though.
- Tkatchenko-Scheffler and exchange-hole dipole model (XDM): Similar in spirit to Grimme's DFT+D, but C_6 coefficients are derived from first principles.

The plane-wave pseudopotential method The atomic arrangement in perfect crystals is described by a periodically repeated *unit cell*. For many interesting physical systems, however, perfect periodicity is absent, but the system is either approximately periodic or periodic in one or two directions or periodic except for a small part. Examples of such systems include surfaces, point defects in crystals, substitutional alloys, heterostructures ("superlattices" and quantum

wells). In all such cases it is convenient to simulate the system with a periodically repeated fictitious *supercell*. The form and the size of the supercell depend on the physical system being studied, e.g.: The study of point defects requires that a defect does not interact with its periodic replica in order to accurately simulate a truly isolated defect. For disordered solids, the supercell must be large enough to guarantee a significant sampling of the configuration space. For surfaces, one uses a crystal slab alternated with a slab of empty space, both large enough to ensure that the bulk behavior is recovered inside the crystal slab and that the surface behavior is unaffected by the presence of the periodic replica of the crystal slab.
In the examples mentioned above, the supercell approach is usually more convenient than the "cluster" one', that is, simulating an extended system by taking a finite piece of material, due to the absence of an abrupt termination in the supercell approach. Even finite systems (molecules, clusters) can be studied using supercells. Enough empty space between the periodic replicas of the finite system must be left so that the interactions between them are weak. The use of supercells for the simulation of molecular or completely aperiodic systems (liquids, amorphous systems) has become quite common in recent years, in connection with first-principles simulations (especially molecular dynamics simulations) using a *plane-wave* (PW) basis set. In fact there are important computational advantages in the use of PW's that may offset the disadvantage of inventing a periodicity where there is none.
The size of the unit cell - the number of atoms and the volume - is very important. Together with the type of atoms it determines the difficulty of the calculation: large unit cells mean large calculations. Unfortunately many interesting physical systems are described, exactly or approximately, by large unit cells.

Plane-wave basis set In the following we will assume that our system is a crystal with lattice vectors $\boldsymbol{R}$ and reciprocal lattice vectors $\boldsymbol{G}$. It is not relevant whether we are dealing with a truly periodic crystal or with a supercell describing an aperiodic system. The KS orbitals are classified by a band index and a Bloch vector $\boldsymbol{k}$ in the Brillouin Zone (BZ).
A PW basis set is defined as $\langle \boldsymbol{r}|\boldsymbol{k}+\boldsymbol{G}\rangle = V^{-½}\, e^{i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}$ for all $\boldsymbol{G}$ vectors such that $\hbar^2|\boldsymbol{k}+\boldsymbol{G}|^2/2m \le E_{cut}$, where V is the crystal volume, E_{cut} is a cutoff on the kinetic energy of PW's (from now on, simply *the cutoff*). PW's have many attractive features: they are simple to use, unbiased (there is no freedom in choosing PW's: the basis is fixed by the crystal structure and by the cutoff), orthonormal, *complete* in the limit of infinite cutoff. Checking for convergence can be easily done by increasing the only relevant convergence parameter: the cutoff.
Unfortunately the extended character of PW's makes it very difficult to accurately reproduce localized functions such as the charge density around a nucleus or even worse, the orthogonalization "wiggles" due to the presence of *core* states in atoms. In order to describe features varying on a length scale δ, one needs Fourier components up to $q \sim 2\pi/\delta$. In a solid, this means a number of PW's $N_{pw} \sim 4\pi\,(2\pi/\delta)^3/3\Omega_{BZ}$ (Ω_{BZ} is the volume of the BZ).

A simple estimate for diamond is instructive. The 1*s* state of the carbon atom has its maximum around 0.3 a.u., so δ=0.1 a.u. is a reasonable value (1 a.u.=1 Bohr radius=0.52977 Å). Diamond has a face-centered cubic (fcc) lattice, lattice parameter a_0=6.74 a.u., $\Omega_{BZ}=(2\pi)^3/(4a_0^3)$. One finds N_{pw}~ 250,000 PW's, clearly too much for practical use.

Pseudopotentials The idea of replacing the full atom with a much simpler *pseudoatom* having valence electrons only arises naturally (apparently in a 1934 paper by Fermi for the first time). *Pseudopotentials*} (PP's) have been widely used in solid-state physics starting from the 1960's. In earlier approaches PP's were devised to reproduce some known experimental solid-state or atomic properties such as energy gaps or ionization potentials. Other types of PP's were obtained from band structure calculations with the OPW, orthogonalized PW, basis set, by separating the smooth (PW) part from the orthogonalization part in the wavefunctions. The first modern version of PP's are the so-called *norm-conserving PPs*. These are *atomic* potentials which build to mimic the scattering properties of the true atom. For a given reference atomic configuration,a norm-conserving PP must fulfill the following conditions:

1. all-electron and pseudo-orbitals must have the same energy, and
2. they must be the same beyond a given "core radius"' r_c, usually located around the outermost maximum of the atomic orbital;
3. the pseudo-charge and the true charge contained in the region $r<r_c$ must be the same.

This last condition explains the name "norm-conserving". There is an historical reason for this: some earlier PP's violated condition 3 (this was known as the "orthogonality hole" problem). Note that the definition "all-electron", here and in the following, refers to a KS calculation that includes core electrons, not to a many-electron wavefunction.

Norm-conserving PP are smooth functions and a relatively small PW basis set is sufficient. They are *nonlocal* because it is usually impossible to mimic the effect of orthogonalization to core states on different angular momenta ℓ with a single function. There is a PP for every ℓ:

$$V^{ps} \equiv V_{\rm loc}(r) + \sum_\ell V_\ell(r)\, P_\ell = V_{\rm loc}(r) + \sum_{\ell m} V_\ell(r) Y_{\ell m}(\boldsymbol{r})\, \delta(r-r')\, Y_{\ell m}^*(\boldsymbol{r}'),$$

where $P_l = |\ell\rangle\langle\ell|$ is the projection operator on states of angular momentum ℓ. The local part of the potential has a long-range tail going like $V_{\rm loc}(r)\cong -Z_v\, e^2/r$, where Z_v is the number of valence electrons. PP's are however seldom used in this form. For computational reasons, they are recast into a fully nonlocal, separable form. The nonlocality of PP's introduces some additional but limited complications in the calculation. In particular, one has to do the following generalization to the DFT formalism: $\int V(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} \rightarrow \sum_i \langle\psi_i|V|\psi_i\rangle = \sum_i \int \psi_i^*(\boldsymbol{r}) V(\boldsymbol{r},\boldsymbol{r}') \psi_i(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}'$

Old and recent experience shows that PP's are equivalent to the *frozen core approximation*: PP and all-electron calculations on the same systems yield almost indistinguishable results (except for those cases in which core states are not sufficiently frozen). It should be remarked that the use of PP's is not limited to PW basis sets: PP's can be (and are actually) used in conjunction with localized basis sets as well.

Ultrasoft PPs and PAW Norm-conserving PP's are still quite "hard" - that is, they contain a significant amount of Fourier components with large q - for a number of atoms, such as N, O, F, and the first row of transition metals. For these atoms little is gained in the pseudization, because there are no orthonormality wiggles that can be removed from the $2p$ and $3d$ states, respectively. *Ultrasoft* PP's have been devised that are much softer than ordinary norm-conserving PP's, at the price of a considerable additional complexity. The heavy formalism of ultrasoft PP's tends to hide the underlying logic (and physics). An alternative approach, called *Projector Augmented Waves* (PAW), is much more transparent. Moreover PAW includes as special cases a number of other methods and provides a simple and consistent way to reconstruct all-electron orbitals from pseudo-orbitals. These are needed for reliable calculation of such observables as NMR chemical shifts and hyperfine coupling coefficients.

The idea of PAW is to find a *mapping* between the all-electron and the pseudo orbitals via a suitable *linear operator*. The pseudo-orbital must be a smooth object that can be expanded into a small PW basis set. Let us consider for simplicity the case of a single atom in the system. In a core region ($r<r_c$) centered around the atom, the mapping is defined as

$$|\phi_i^{ae}\rangle = (1 + T)\,|\phi_i^{ps}\rangle$$

where the functions $|\phi_i^{ae}\rangle$ are solutions, regular at the origin but not necessarily bound, of the all-electron atomic KS equation; the functions $|\phi_i^{ps}\rangle$ are the corresponding pseudo-functions, that are much smoother in the core region and join smoothly to the $|\phi_i^{ae}\rangle$ at the border of core region ($r=r_c$). Outside the core region, we set T=0; inside, we assume that we may write a pseudo-orbital $|\psi^{ps}\rangle$ for our molecular or solid-state system as a sum over the atomic pseudo-waves $|\phi_i^{ps}\rangle$: $|\psi^{ps}\rangle = \sum_i c_i\,|\phi_i^{ps}\rangle$. By applying the operator (1 + T) to both sides of the above expansion we find $|\psi^{ae}\rangle = \sum_i c_i\,|\phi_i^{ae}\rangle$, where $|\psi^{ae}\rangle$ is the all-electron wavefunction. The above result can be recast into the form

$$|\psi^{ae}\rangle = |\psi^{ps}\rangle + \sum_i c_i\,(|\phi_i^{ae}\rangle - |\phi_i^{ps}\rangle)$$

It remains to define the c_i coefficients. Let us introduce the projectors β_i such that

$$\langle\beta_i|\,\phi_j^{ps}\rangle = \delta_{lm}, \qquad \sum_i |\phi_i^{ps}\rangle\langle\beta_i| = I.$$

It is easy to verify that $c_i = \langle\beta_i|\,\psi^{ps}\rangle$ and that we can write $|\psi^{ae}\rangle = |\psi^{ps}\rangle + \sum_i (|\phi_i^{ae}\rangle - |\phi_i^{ps}\rangle)\langle\beta_i|\,\psi^{ps}\rangle$, or $|\psi^{ae}\rangle = (1+T)\,|\psi^{ps}\rangle$. The operator $T=\sum_i (|\phi_i^{ae}\rangle - |\phi_i^{ps}\rangle)\langle\beta_i|$ thus projects out the atomic pseudo-states ϕ_i^{ps} and replaces them with the all-electron states ϕ_i^{ae}. The T operator is a purely atomic quantity that is obtained from a judicious choice of the ϕ_i^{ae} all-electron atomic states, the corresponding pseudo-states ϕ_i^{ps}, and the projectors β_i. The equations to solve in the PAW method are then obtained by inserting the above form for ψ^{ae} in the energy functional and by finding its minimum with respect to the variation of the smooth part only, ψ^{ps}. An important feature of the PAW method is that the charge density is no longer given simply by the square of the orbitals, but it contains in general an additional (*augmentation*) term:

$$n(\boldsymbol{r}) = \sum_i |\psi_i^{ps}(\boldsymbol{r})|^2 + \sum_i \sum_{\ell m} \langle\psi_i^{ps}|\,\beta_\ell\rangle\; q_{\ell m}\,(\boldsymbol{r})\,\langle\beta_m|\psi_i^{ps}\rangle$$

where $q_{\ell m}(\boldsymbol{r}) = \phi_\ell^{ae}(\boldsymbol{r})\,\phi_m^{ae}(\boldsymbol{r}) - \phi_\ell^{ps}(\boldsymbol{r})\,\phi_m^{ps}(\boldsymbol{r})$ (using the completeness relation). Conversely, the pseudo-orbitals are no longer orthonormal, but obey a generalized orthonormality relation:

$$\langle \psi_i^{ps} | S | \psi_j^{ps} \rangle = \delta_{ij}, S = I + \sum_{\ell m} | \beta_\ell \rangle\, Q_{\ell m} \langle \beta_m |, \quad Q_{\ell m} = \int q_{\ell m}(\boldsymbol{r}) d\boldsymbol{r}.$$

In the PAW formalism, the $q_{\ell m}(\boldsymbol{r})$ functions are stored on an auxiliary radial grid centered on atoms. Ultrasoft PP's can be derived from PAW by *pseudizing* the $q_{\ell m}(\boldsymbol{r})$, that can be thus expanded into plane waves. Norm-conserving PP's in the separable form can be derived from PAW if the atomic states $\phi_\ell^{ae}(\boldsymbol{r})$ and $\phi_\ell^{ps}(\boldsymbol{r})$ obey the norm-conservation rule (thus S=1).

Brillouin-Zone sampling In order to calculate the charge density $n(\boldsymbol{r})$ in a periodic systems, one has to sum over an infinite number of $\boldsymbol{k}$-points:

$$n(\boldsymbol{r}) = \sum_{\boldsymbol{k}i} |\psi_{\boldsymbol{k}i}(\boldsymbol{r})|^2$$

where the index i runs over occupied bands. Assuming periodic (Born-Von Kàrmàn) boundary conditions: $\psi(\boldsymbol{r}+L_1\boldsymbol{R}_1)=\psi(\boldsymbol{r}+L_2\boldsymbol{R}_2)=\psi(\boldsymbol{r}+L_3\boldsymbol{R}_3)=\psi(\boldsymbol{r})$, a crystal has $L=L_1L_2L_3$ allowed $\boldsymbol{k}$-points (L is also the number of unit cells). In the *thermodynamic* limit of an infinite crystal, $L\to\infty$, the discrete sum over $\boldsymbol{k}$ becomes an integral over the BZ. Experience shows that this integral can be approximated by a discrete sum over an affordable number of $\boldsymbol{k}$-points, at least in insulators and semiconductors. When present, symmetry can be used to further reduce the number of calculations to be performed. Only one $\boldsymbol{k}$-point is left to represent each *star* - the set of $\boldsymbol{k}$-points that are equivalent by symmetry -with a weight w that is proportional to the number of $\boldsymbol{k}$-points in the star. The infinite sum over the BZ is replaced by a discrete sum over a set of points $\boldsymbol{k}_n$ and weights w_n:

$$(1/L)\ \sum_{\boldsymbol{k}} f_{\boldsymbol{k}}(\boldsymbol{r}) \to \sum_n w_n f_{\boldsymbol{k}_n}(\boldsymbol{r}).$$

The resulting sum is then symmetrized to get the charge density. Suitable sets for BZ sampling in insulators and semiconductors are called *special points*. This name is somewhat misleading: in most cases those sets just form uniform grids in the BZ.
In metals things are more difficult because one needs an accurate sampling of the Fermi surface. The DFT extension to fractionary occupation numbers is used. The *smearing* and the *tetrahedron* techniques, or variations of the above, are generally used.
In large units cells and supercells, the $\boldsymbol{k}$-point grid is often limited to the Γ point ($\boldsymbol{k}$=0).

Iteration to self-consistency The KS equations must be solved *self-consistently*. We supply an input charge density $n_{\text{in}}(\boldsymbol{r})$ to the KS equations and we get an output charge density $n_{\text{out}}(\boldsymbol{r})$. Such procedure defines a functional: $n_{\text{out}}(\boldsymbol{r}) = A[n_{\text{in}}(\boldsymbol{r})]$. At self-consistency, $n(\boldsymbol{r}) = A[n(\boldsymbol{r})]$. The first algorithm that comes to the mind is to simply use $n_{\text{out}}(\boldsymbol{r})$ as the new input charge density, setting $n^{(i+1)}{}_{in}(\boldsymbol{r}) = n^{(i)}{}_{\text{out}}(\boldsymbol{r})$, where the superscripts indicate the iteration number. Unfortunately this usually *does not work*. The reason is that there is no guarantee that the error on output is smaller than the error on input. If $\delta n_{in}(\boldsymbol{r})$ is the error on input, the error on output, close to self-consistency, will be $\delta n_{out}(\boldsymbol{r}) \cong \int (\delta A[n(\boldsymbol{r})]/\delta n(\boldsymbol{r}'))\delta n_{in}(\boldsymbol{r}')\, d\boldsymbol{r}' \equiv (J\delta n_{in})(\boldsymbol{r})$, which may or may not be

smaller than the input error: it depends on the size of the largest eigenvalue, e_J, of the operator J, which is related to the dielectric response of the system. Usually $e_J>1$ and the iteration does not converge.

A simple algorithm that generally works, although sometimes slowly, is *simple mixing*. A new input charge density is generated by mixing the input and output charges:

$$n^{(i+1)}{}_{in}(\boldsymbol{r}) = (1-\alpha)n^{(i)}{}_{in}(\boldsymbol{r}) + \alpha n^{(i)}{}_{\text{out}}(\boldsymbol{r})$$

The value of α must be chosen empirically in order to get fast convergence. The error with respect to self-consistency becomes $\delta n_{out} \cong [(1-\alpha)+\alpha J]\delta n_{in}$ and it is easily seen that the iteration converges if $\alpha<|1/e_J|$. In general, the convergence is easier for small cells and symmetric systems, more difficult for larger cells, low symmetry, cells elongated along one directions, surfaces. Relatively big values (α=0.3÷0.5) can be chosen in "easy"' systems, smaller values are appropriate for cases of difficult convergence.

Better results are obtained with more sophisticated algorithms (to name a few: Anderson, Broyden, Direct Iteration in Inverse Space, DIIS) that use informations collected from several preceding iterations. Let us sketch the logic of such algorithms. We have a sequence of $n^{(i)}{}_{in}$ producing $n^{(i)}{}_{out}$ from preceding iterations. We look for the linear combination of input $n^{(i)}{}_{new}$:

$$n^{(i)}{}_{new} = \sum_j c_j \; n^{(j)}{}_{in}, \;\; \sum_j c_j = 1$$

that minimizes an appropriate norm $||n^{(i,out)}{}_{new} - n^{(j,in)}{}_{new}||$. Close to self-consistency, one can write $||n^{(i,out)}{}_{new} - n^{(j,in)}{}_{new}|| \cong || \sum_j c_j \, (n^{(j)}{}_{out} - n^{(j)}{}_{in}) \, ||$ and the coefficients c_j are determined by imposing that such norm is minimum. Then we mix $n^{(j,in)}{}_{new}$ with $n^{(i,out)}{}_{new} = \sum_j c_j \; n^{(j)}{}_{out}$ (using e.g. simple mixing) to obtain $n^{(i+1)}{}_{in} = (1-\alpha)\sum_j c_j \; n^{(j)}{}_{in} + \alpha \sum_j c_j \; n^{(j)}{}_{out}$.

Diagonalization of the Hamiltonian When the KS orbitals are expanded on a finite basis set, the KS equations take the form of a secular equation: $\sum_{G'} H(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}')\psi_{k,i}(\boldsymbol{G}') = \varepsilon_{k,i}\psi_{k,i}(\boldsymbol{G})$, where the matrix elements of the Hamiltonian have the form

$$H(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}') = (\hbar^2/2m)|\boldsymbol{k}+\boldsymbol{G}|^2 \delta_{G-G'} + V_{scf}(\boldsymbol{G}-\boldsymbol{G}') + V_{loc}(\boldsymbol{G}-\boldsymbol{G}') + V_{nl}(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}')$$

The term $V_{scf}(\boldsymbol{G}-\boldsymbol{G}')$ is the Fourier transform of the XC and Hartree potential:

$$V_{scf}(\boldsymbol{G}-\boldsymbol{G}') = (1/\Omega) \int V_{scf}(\boldsymbol{r}) \, e^{i(\boldsymbol{G}-\boldsymbol{G}')\boldsymbol{r}} d\boldsymbol{r}$$

(Ω is the volume of the unit cell, the integration is on a single cell) and the same applies to V_{loc} that comes from the local term in the PP's. The nonlocal contribution V_{nl} comes from PP's:

$$V_{NL}(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}') = (1/\Omega) \int V_{NL}(\boldsymbol{r},\boldsymbol{r}') \, e^{i(\boldsymbol{k}+\boldsymbol{G})\boldsymbol{r}} d\boldsymbol{r} \, e^{i(\boldsymbol{k}+\boldsymbol{G}')\boldsymbol{r}'} d\boldsymbol{r} \, d\boldsymbol{r}'$$

With USPP and PAW, the secular problem generalizes to

$$\sum_{G'} H(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}')\psi_{\boldsymbol{k},i}(\boldsymbol{G}') = \varepsilon_{\boldsymbol{k},i}\sum_{G'} S(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}') \, \psi_{\boldsymbol{k},i}(\boldsymbol{G}),$$

where S is the overlap matrix. We need ot find the lowest N_b eigenvalues and eigenvectors (only valence states for insulators, a few more for metals) of an N_{pw} x N_{pw} Hermitian matrix. This task is performed using *iterative* techniques that do not require to store the entire matrix but just need products $H\psi$. These are computed quickly and effectively using the *dual-space technique*, jumping from real to reciprocal space and back using FFT, Fast Fourier-Transform.

Bibliography

1. *Electronic Structure: Basic Theory and Practical Methods*, R. M. Martin, Cambridge University Press, 2012. An excellent introductory book by one of the leading scientists in this field.
2. *Materials Modelling using Density Functional Theory: Properties and Predictions,* F. Giustino, Oxford University Press, 2014. Another excellent introductory book.
3. *Density Functional Theory: A Practical Introduction*, D. Scholl and J.A. Steckel, Wiley (2009). A recent introductory book on Density-Functional Theory
4. *Density Functional Theory*, R.M. Dreizler and E.K.U. Gross, Springer-Verlag, Berlin (1990). Less recent and very formal, but still "the" reference if you want to *really* understand DFT.

Extremely localized molecular orbitals in quantum crystalography

Alessandro Genoni
CNRS & University of Lorraine, Metz, France
alessandro.genoni@univ-lorraine.fr

1. Introduction

As is well known, chemistry is still usually and largely interpreted in terms of local concepts, at least at a first-approximation level. For instance, when looking at molecules, the very first interpretation of their electronic structure is still based on the corresponding Lewis diagrams, which show chemical bonds as simple lines drawn between atoms. Furthermore, in organic chemistry, the reasoning is always strongly based on functional groups, which are small molecular fragments that keep their main features in different compounds, and reaction mechanisms are usually depicted by means of arrows that represent electron-flows occurring during the breaking and formation of bonds.

Nevertheless, in theoretical chemistry, excluding the Valence-Bond (VB) approaches,[1,2] which have been developed only by a restricted group of scientists and can be applied only to small systems due their large computational cost, the most widely used MO-based methods[3] (namely, techniques based on Molecular Orbitals) provide a completely delocalized picture of molecular electronic structure. This is the reason why, over the years, several computational strategies have been successfully developed in order to recover traditional chemical concepts from quantum chemistry calculations. In this context, it is necessary to distinguish between *a posteriori* and *a priori* techniques.

Among the former, a prominent role is obviously occupied by the so-called topological approaches. The most popular one is the Quantum Theory of Atoms in Molecules[4] (QTAIM) that primarily enables to rationalize theoretical (but also experimental) electron densities in terms of nuclear attractors, bond paths and bond, ring & cage critical points, all of them descriptors allowing to establish a link between quantum mechanical calculations and the traditional chemical perception. Other important *a posteriori* topological methods are also the ones that fully exploit the information content of the one-electron density matrices obtained from quantum chemistry calculations and that provide real-space functions strictly related to the electron pairing, such as the Electron Localization Function (ELF),[5] the Electron Localizability Indicator (ELI)[5] and the Localized Orbital Locator (LOL).[7]

Of course, in the framework of the *a posteriori* techniques, it is also worth mentioning the traditional methods of quantum chemistry for the localization of the completely delocalized canonical Hartree-Fock Molecular Orbitals (MOs, see Figure 1A) obtained as solutions of the canonical Hartree-Fock equations.[3] In fact, these strategies consist in performing unitary

transformations of the canonical Hartree-Fock MOs, thus providing an equivalent set of Molecular Orbitals that maximize or minimize a chemically/physically meaningful functional. In this context, some of the most representative methods are: i) the Foster-Boys approach,[8,9] which aims at minimizing the spatial extension of the Molecular Orbitals, ii) the Edminston-Ruedenberg technique,[10,11] whose goal is to maximize the "auto-repulsion energy", iii) the Von Niessen strategy,[12] which tries to maximize the "charge density overlap", and iv) the Pipek-Mezey procedure,[13,14] which maximizes a functional based on the Mulliken population. All the methods mentioned above provide Molecular Orbitals that are mainly localized on small molecular units, but that are also characterized by the so-called "orthogonalization tails" that extend beyond the main localization regions (see Figure 1B). For this reason, they cannot be unambiguously associated with small molecular fragments. In order to obtain Molecular Orbitals strictly localized on (and, consequently, unambiguously associable with) small molecular units (e.g., atoms, bonds or functional groups), it is actually necessary to resort to the Extremely Localized Molecular Orbitals[15,16] (ELMOs, see Figure 1C), which are the main topic of this lecture.

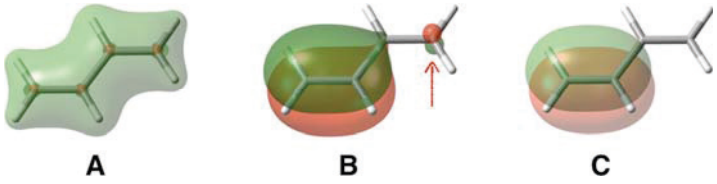


Figure 1. (A) Completely delocalized canonical Hartree-Fock Molecular Orbital, (B) Pipek-Mezey Localized Molecular Orbital with typical orthogonalization tail and (C) Extremely Localized Molecular Orbital for a C-C bond of butadiene.

The technique to obtain ELMOs has been introduced by Stoll and coworkers in 1980[15] and is strictly connected both to the earlier "group function method" introduced by McWeeny in the 1960s[17-19] and to many other theoretical approaches that have been developed over the years[20-29] to decompose global electronic wavefunctions into functions describing smaller subsets of electrons. The Stoll method[15] is a typical *a priori* strategy of theoretical chemistry that introduces traditional chemical concepts before performing the calculations. In fact, the technique is based on the *a priori* definition of a localization scheme that subdivides the investigated chemical compound into different subunits according to the chemical intuition (see Figure 2A). Of course, the most intuitive (and most adopted) choice is a localization pattern corresponding to the Lewis structure of the system under exam, which is therefore characterized by atomic fragments for the core and lone-pair electrons and by bond fragments for the bond electron-pairs. As a consequence of this preliminary fragmentation, the Molecular Orbitals corresponding to the different subunits are expanded on local basis-sets (i.e., sets of basis functions (or atomic orbitals) centred only on the atoms belonging to the fragments) and the ELMOs are afterwards obtained by variationally minimizing the energy associated with the single Slater determinant constructed with them (from now on, ELMO wavefunction). It is

possible to show that this is equivalent to solving these modified Hartree-Fock equations for each fragment:[15]

$$\hat{F}^j \left| \varphi_\alpha^j \right\rangle = \varepsilon_\alpha^j \left| \varphi_\alpha^j \right\rangle \qquad (1)$$

with $\hat{F}^j$ as the modified Fock operator for the j-th fragment given by

$$\hat{F}^j = \left(1 - \hat{\rho} + \hat{\rho}^{j\dagger}\right) \hat{F} \left(1 - \hat{\rho} + \hat{\rho}^j\right) \qquad (2)$$

where $\hat{F}$ is the usual Fock operator of quantum chemistry, $\hat{\rho}^j$ is the local density operator for the j-th subunit, which depends only on the occupied ELMOs of the fragment, and $\hat{\rho}$ is the global density operator, which, on the contrary, depends on all the occupied ELMOs of the system and, for this reason, couples all the modified Hartree-Fock equations associated with the different subunits.

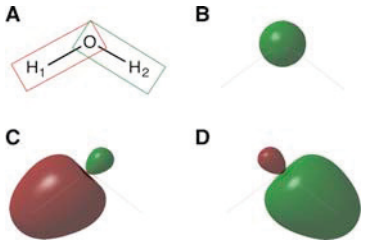


Figure 2. (A) Localization scheme for the water molecule with the two overlapping bond fragments O–H1 and O–H2 explicitly shown. For the sake of clarity, the atomic fragment O, which describes core and lone-pair electrons of the oxygen atom, is not depicted; (B) ELMO for the core electrons of the oxygen atom; (C) ELMO for the O-H1 bond; (D) ELMO for the O-H2 bond.

Since both atomic and bond fragments are simultaneously considered in the calculations, the Stoll method provides MOs strictly localized on both atomic (see Figure 2B) and bond (see Figures 2C and 2D) subunits, thus taking into account chemical bonding directly. Due to their extreme localization, ELMOs can be indeed associated with small molecular units without ambiguity and, following a well-defined rotation procedure,[30,31] they can be transferred from molecule to molecule as electronic LEGO building blocks[31,32] in order to reconstruct wavefunctions and electron densities of macromolecules (e.g. proteins) almost instantaneously (see Figure 3).

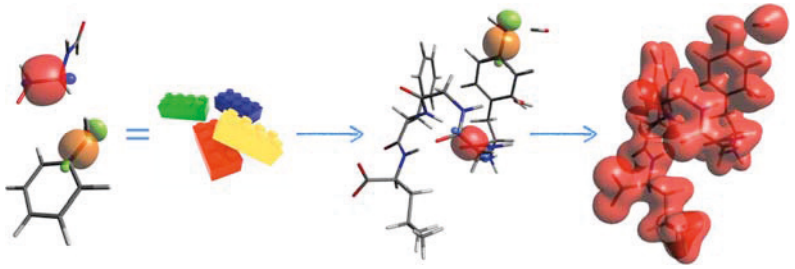


Figure 3. Extremely Localized Molecular Orbitals as elementary, electronic LEGO building blocks to reconstruct wavefunctions and electron densities of macromolecules.

Other than introducing the concept of ELMOs, this lecture will also consist in discussing the current use of ELMOs in quantum crystallography. In, particular, it will be shown how their strict localization has been exploited i) to introduce traditional chemical concepts *a priori* within the context of the X-ray constrained wavefunction (XCW) approach[33-40] (see section 2) and ii) to construct new libraries of Extremely Localized Molecular Orbitals[41] with the final goal of reconstructing electron densities and refining crystallographic structures of proteins (see section 3). Concerning this last point, preliminary results obtained through the coupling of the recently constructed ELMO-libraries[41] with the Hirshfeld Atom Refinement (HAR)[42-45] will be also briefly discussed.

2. ELMOs in the framework of the X-ray constrained wavefunction approach

As just mentioned above, in this section it will be shown how the strict localization of the Extremely Localized Molecular Orbitals has been exploited to introduce *a priori* the traditional chemical picture of molecular electronic structure within the X-ray constrained wavefunction approach[33-40]. In particular, in section 2A the main topic will be the X-ray constrained ELMO (XC-ELMO) method,[46-49] a technique that combines the XCW strategy introduced by Jayatilaka with the approach proposed by Stoll to obtain Extremely Localized Molecular Orbitals. In section 2B, the focus will be on the more recent X-ray constrained ELMO-Valence Bond[50,51] (XC-ELMO-VB) method, another computational technique that exploits the strict localization of the ELMOs to keep a high chemical interpretability of the results and that can be seen as the first-"prototype" many-determinant XCW approach.

2.A The X-ray constrained ELMO method

Since the XC-ELMO technique is a combination of the original XCW approach[33-40] with the traditional ELMO method,[15] its working equations can be simply obtained by minimizing the energy associated with the ELMO wavefunction (as in the Stoll method[15]) with the additional constraint of reproducing a set of collected structure factor amplitudes $\{|F_{\mathbf{h}}^{obs}|\}$within a predetermined desired agreement (in analogy with the Jayatilaka technique[33-40]). In other words, it is necessary to look for those Extremely Localized Molecular Orbitals that minimize the following functional:

$$J_{ELMO}[\boldsymbol{\varphi}] = E_{ELMO}[\boldsymbol{\varphi}] + \lambda_J (\chi^2[\boldsymbol{\varphi}] - \Delta) \quad (3)$$

where E_{ELMO} is the energy associated with the ELMO wavefunction for the reference crystal unit, λ_J is an external adjustable parameter that is varied during the calculations and that represents the strength of the external constraint, χ^2 is the measure of the statistical agreement between calculated and experimental structure factor amplitudes, Δ is the desired agreement (typically fixed to 1.0), and $[\boldsymbol{\varphi}]$ stresses the functional dependence on the occupied ELMOs. In particular, χ^2 is expressed like this:

$$\chi^2 = \frac{1}{N_r - N_p} \sum_{\mathbf{h}} \frac{\left(\eta|F_{\mathbf{h}}^{calc}| - |F_{\mathbf{h}}^{obs}|\right)^2}{\sigma_{\mathbf{h}}^2} \quad (4)$$

with N_r as the number of collected X-ray diffraction data, N_p as the number of adjustable parameters (in this case only the external multiplier λ_j), $\boldsymbol{h}$ as the triad of Miller indices labeling the reflection, $\sigma_{\boldsymbol{h}}$ as the standard uncertainty corresponding to each observed structure factor amplitude $|F_{\boldsymbol{h}}^{obs}|$, and η as a scale factor that is properly determined in order to minimize χ^2. Through a quite simple mathematical derivation[46,47] it was possible to show that the minimization of functional (3) is equivalent to solving self-consistently a new set of equations that are substantially the equations proposed by Stoll *et al.* for the original ELMO method[15] (see equation (1)), but with the modified Fock operators for the different fragments containing two additional terms that take into account the effect of the experimental data:[46,47]

$$\begin{aligned}\hat{F}^{j,exp} &= \left(1-\hat{\rho}+\hat{\rho}^{j\dagger}\right)\hat{F}\left(1-\hat{\rho}+\hat{\rho}^{j}\right)+ \\ &+ \lambda\sum_{\boldsymbol{h}} K_{\boldsymbol{h}}\, Re\{F_{\boldsymbol{h},calc}\}\left(1-\hat{\rho}+\hat{\rho}^{j\dagger}\right)\hat{I}_{\boldsymbol{h},R}\left(1-\hat{\rho}+\hat{\rho}^{j}\right)+ \\ &+ \lambda\sum_{\boldsymbol{h}} K_{\boldsymbol{h}}\, Im\{F_{\boldsymbol{h},calc}\}\left(1-\hat{\rho}+\hat{\rho}^{j\dagger}\right)\hat{I}_{\boldsymbol{h},C}\left(1-\hat{\rho}+\hat{\rho}^{j}\right) \qquad (5)\end{aligned}$$

with

$$K_{\boldsymbol{h}} = \frac{2\eta}{N_r - N_p}\,\frac{\eta F_{\boldsymbol{h},calc} - F_{\boldsymbol{h},obs}}{\sigma^2_{\boldsymbol{h},obs}\; F_{\boldsymbol{h},calc}} \qquad (6)$$

and $\hat{I}_{\boldsymbol{h},R}$ and $\hat{I}_{\boldsymbol{h},C}$ respectively as real and imaginary parts of the structure factor operator $\hat{I}_{\boldsymbol{h}}$ defined like this:

$$\hat{I}_{\boldsymbol{h}} = \sum_{k=1}^{N_m} e^{i2\pi(\boldsymbol{R}_k\boldsymbol{r}+\boldsymbol{r}_k)\cdot(\boldsymbol{B}\boldsymbol{h})} = \hat{I}_{\boldsymbol{h},R} + i\,\hat{I}_{\boldsymbol{h},C} \qquad (7)$$

where $\boldsymbol{B}$ is the reciprocal lattice matrix and N_m is the number of unit-cell equivalent positions, which are related to the reference one through the crystal symmetry operations $(\boldsymbol{R}_k, \boldsymbol{r}_k)$. The strategy has been properly tested by performing calculations that exploited both high- and low-quality crystallographic data and that used sets of atomic basis functions of different size and qualities.[46-48] It mainly emerged that the determination of X-ray constrained ELMOs is really straightforward. Nevertheless, as already observed for the original XCW technique, good statistical agreements with the experimental X-ray diffraction measurements are reached provided that high-quality crystallographic data and sufficiently flexible basis-sets are used. From the preliminary tests it was also observed that the X-ray constrained procedure entails significant redistributions of the electronic charges both in the valence and in the core regions. In a more recent study,[49] the capabilities of the new XC-ELMO approach have been also assessed by comparing the results of corresponding X-ray constrained Hartree-Fock (XC-HF) and X-ray constrained ELMO calculations. This enabled to study in detail the effects of introducing an *a priori* localization of Molecular Orbitals in the framework of Jayatilaka's XCW approach. It was observed that, if small basis-sets are used in the calculations, the MOs localization has indeed a strong influence on the final electron density. Nevertheless, employing larger sets of basis functions, the XC-HF and the XC-ELMO techniques tend to provide more similar results. This

can be explained considering that the flexibility of larger basis-sets allows to increasingly exploit the information provided by the experimental data and, consequently, to gradually mitigate the initial bias imposed on the electronic structure by a pre-defined localization. Therefore, when adequate basis-sets are used, the introduction of a localization scheme represents a less severe approximation and, at the same time, it enables to directly include traditional chemical concepts in the X-ray constrained wavefunctions.

2.B The X-ray constrained ELMO-Valence Bond method

All the current versions of the XCW approach are based on a single Slater determinant wavefunction *ansatz.* However, in order to determine the weights of different resonance structures in molecular systems characterized by a multi-reference character, it has been recently developed the X-ray constrained ELMO-Valence Bond (XC-ELMO-VB) method.[50] Unlike the usual X-ray constrained wavefunction strategies, in the XC-ELMO-VB technique each crystal unit wavefunction is written in the following form:

$$| \Psi_{XC-ELMO-VB} \rangle = \sum_i C_i \, | \Psi_i \rangle \quad (8)$$

where the functions $\{| \Psi_i \rangle\}$ are single Slater determinants that describe all the possible resonance structures of the system in exam. In this case, they consist in ELMO wavefunctions that are pre-determined by means of unconstrained ELMO calculations based on localization schemes corresponding to the different resonance structures. For example, if one were interested in studying the benzene molecule, the wavefunction $| \Psi_{XC-ELMO-VB} \rangle$ would be a linear combination i) of the ELMO wavefunction corresponding to the localization scheme for resonance structure A and ii) of the ELMO wavefunction associated with the localization scheme for resonance structure B (see Figure 4).

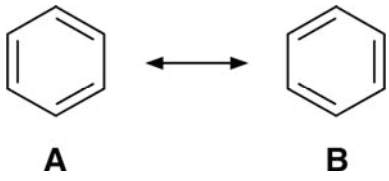


Figure 4. Resonance structures of benzene

In the current version of the method,[50,51] the pre-optimized ELMOs are kept frozen, while the coefficients $\{C_i\}$ in equation (8) are determined by minimizing the following functional:

$$J[\boldsymbol{C}] = E[\boldsymbol{C}] + \lambda_J \, (\chi^2[\boldsymbol{C}] - \Delta) \quad (9)$$

where E is the energy of the system associated with wavefunction (8), χ^2 and Δ have the same meaning seen in equation (3) and $\boldsymbol{C}$ indicates the functional dependence of J, E and χ^2 on the coefficient $\{C_i\}$ of expansion (8).

For the sake of completeness, it is important to note that, due to the non-orthogonality of the

Slater determinants $\{|\Psi_i\rangle\}$, the coefficients $\{C_i\}$ do not immediately provide the real weights associated with the corresponding resonance structures. These weights are actually given by the Chirgwin-Coulson coefficients[52] defined as:

$$K_i = |C_i|^2 + \sum_{j \neq i} C_i \, C_j \, S_{ij} \quad (10)$$

with $S_{ij} = \langle \Psi_i | \Psi_j \rangle$ as the overlap between the ELMO wavefunctions $|\Psi_i\rangle$ and $|\Psi_j\rangle$.

Preliminary test calculations have been initially performed to determine the weights of the resonance structures of naphthalene at different temperatures by exploiting experimental high-resolution X-ray diffraction data.[50] The obtained results have interestingly shown that the explicit consideration of experimental structure factors in the determination of the resonance structure weights may lead to results different from those resulting only from the simple energy minimization.

Moreover, the novel XC-ELMO-VB strategy has been more interestingly applied[51] to further interpret the results of a recent charge density study on the syn-1,6:8-13-biscabonyl[14]annulene (BCA) molecule (see Figure 5) conducted by Macchi and coworkers.[53] By performing X-ray diffraction measurements at ambient and high pressures, they have observed a partial suppression of aromaticity when pressure increases.

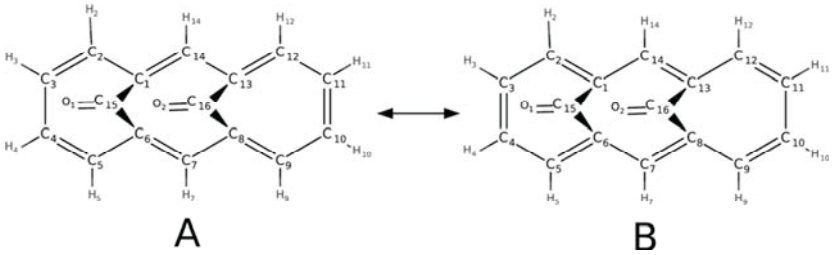


Figure 5. Resonance structures of BCA.

By means of XC-ELMO-VB calculations that exploited crystallographic data (i.e., atomic positions, Anisotropic Displacement Parameters (ADPs) and structure factor amplitudes) collected at different pressures, it has been observed that, while at ambient pressure resonance structures A and B of BCA (see Figure 5) are almost equivalent (i.e., they have almost the same Chirgwin-Coulson coefficients), at high pressures, resonance structure A becomes predominant,[51] confirming the partial rupture of aromaticity experimentally observed by Casati *et al.*[53]

Notwithstanding the encouraging preliminary results, the XC-ELMO-VB strategy represents only the starting point for the development of a new many-determinant X-ray constrained wavefunction method. The next step will consist in also optimizing the ELMO wavefunctions $\{|\Psi_i\rangle\}$ in expansion (8) through the minimization of functional J.

3. ELMO-libraries to refine crystallographic structures of macromolecules

As already mentioned in the introduction, due to their strict localization, the ELMOs can be also exploited as elementary, electronic LEGO-building blocks that can be transferred from molecule to molecule.

In this context, after rigorous preliminary tests on the ELMOs transferability,[31,32] ELMO-libraries that cover all the possible fragments of the twenty natural amino acids in all their possible protonation states have been recently constructed.[41] The ELMOs in the databanks are generally localized on one-atom and two-atom subunits. However, to properly describe situations in which the delocalized nature of the electronic structure is important (e.g., peptide bonds, aromatic rings, carboxylate groups), ELMOs localized on proper three-atom subunits have also been determined and stored (see Figure 6).

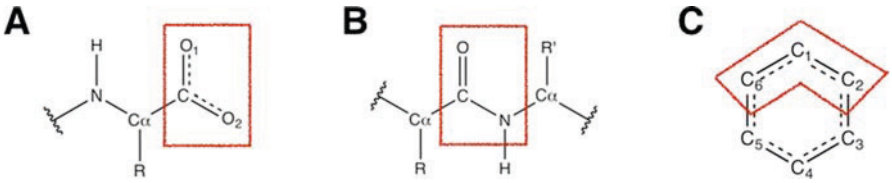


Figure 6. Three-atom fragments used to describe (A) σ and π electrons of the carboxylate groups, (B) the electrons involved in the peptide bonds and (C) the delocalized π electron pairs of the aromatic rings.

To make the ELMO transfer from the libraries to the target structures more and more automatic and fast, a new efficient program has been written. It allows to instantaneously obtain approximate wavefunctions and/or electron densities of very large macromolecules and represents the real starting point for future developments of novel linear scaling ELMO-based strategies aiming at refining high-resolution crystallographic structures of polypeptides and, above all, proteins (see below the very recent coupling of the ELMO-libraries with the Hirshfeld Atom Refinement).

Preliminary test calculations exploiting the ELMO-libraries and the associated program have been already carried out. At first the focus was on the scalability of the ELMOs transfer and, by considering sets of poly-glycines and poly-tryptophans of increasing size, it has been observed that the transfer of ELMOs from the databanks to the target systems is really instantaneous. Of course, the new libraries have been already applied to real proteins, as the antifreeze protein RD1 (PDB code: 1UCS, 64 residues and 997 atoms) and the PEX14 N-terminal domain (PDB code: 5L87, 62 residues and 1019 atoms). For all the considered systems, the new ELMO-libraries and the associated program have allowed to successfully and rapidly compute (e.g., 193.71 seconds for the antifreeze protein RD1) the corresponding electron densities and electrostatic potentials (see Figure 7).

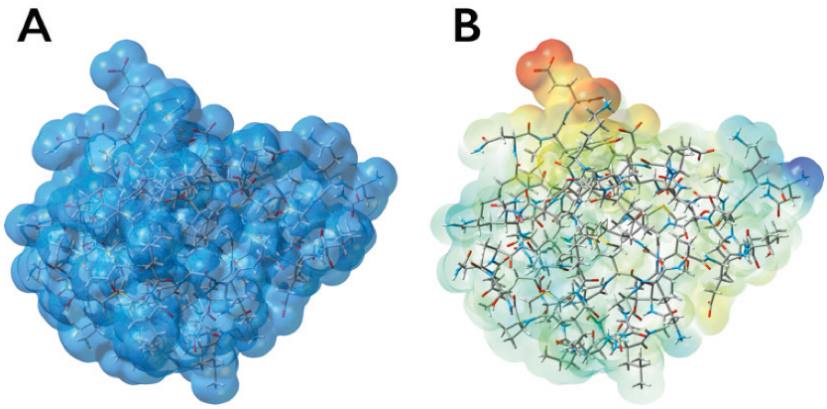


Figure 7. (A) Electron density (0.001 e/bohr^3 isosurface) and (B) electrostatic potential (plotted on the 0.001 e/bohr^3 electron density isosurface) for the antifreeze protein RD1, both of them obtained from the transfer of Extremely Localized Molecular Orbitals from the recently constructed ELMO-libraries (basis-set 6-31G).

Considering the reliable transferability of the ELMOs and the efficiency (in terms of CPU time) with which the new libraries of Extremely Localized Molecular Orbitals are able to reconstruct electron densities of very large systems, the possibility of using the new databanks to refine protein crystallographic structures has been already explored. In particular, the ELMO-databases have been recently coupled with the promising Hirshfeld Atom Refinement (HAR),[42-45] thus extending the applicability of the latter also to macromolecules.[54]

In fact, HAR is a technique that requires a tailor-made quantum mechanical calculation at each step of the refinement and its computational cost unavoidably increases with the size of the systems under exam. It is thus clear that, in its original version, HAR cannot be straightforwardly extended to large molecules. The only way to do that is to couple it with linear-scaling methods of quantum chemistry and this is the reason why HAR has been coupled with the new ELMO-libraries. Although, at the moment, preliminary refinements have been carried out only on crystals of small systems (e.g., dipeptide Gly-Ala), it has been already observed that the new HAR-ELMO technique provides results practically identical to those obtained by means of the traditional HAR method.[54]

4. Conclusions

This lecture is a general overview on the concept of Extremely Localized Molecular Orbitals and on their current use in quantum crystallography. The main message is that, due to their extreme localization and their intrinsic closeness to the traditional chemical perception, ELMOs can be properly exploited i) both to enhance the chemical interpretability of the obtained X-ray constrained wavefunctions (see XC-ELMO and XC-ELMO-VB methods discussed in section 2) ii) and to develop new quantum mechanics-based linear-scaling techniques to refine crystallographic structure of macromolecules (see the ELMO-libraries and the new HAR-ELMO method discussed in section 3).

5. References

1. A. Shurki, *Theor. Chem. Acc.* **116**, 253-261 (2006).
2. P. C. Hiberty and S. Shaik, *J. Comput. Chem.* **28**, 137-151 (2007).
3. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry. Introduction to Advanced Electronic Structure Theory*, Dover Publications, Mineola, New York, USA (1996).
4. R. F. W. Bader, *Atoms in Molecules: a Quantum Theory,* Clarendon Press, Oxford, UK (1990).
5. A. D. Becke and K. E. Edgecombe, *J. Chem. Phys.* **92**, 5397-5403 (1990).
6. M. Kohout, *Int. J. Quantum Chem.* **97**, 651-658 (2004).
7. H. L. Schmider and A. D. Becke, *J. Mol. Struct. (Theochem)* **527**, 51-61 (2000).
8. S. F. Boys, *Rev. Mod. Phys.* **32**, 296-299 (1960)
9. J. M. Foster and S. F. Boys, *Rev. Mod. Phys.* **32**, 300-302 (1960).
10. C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457-464 (1963).
11. C. Edmiston and K. Ruedenberg, *J. Chem. Phys.* **43**, S97-S116 (1965).
12. W. von Niessen, *J. Chem. Phys.* **56**, 4290-4297 (1972).
13. J. Pipek and P. G. Mezey, *Int. J. Quantum Chem.* **34**, 1-13 (1988).
14. J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916-4926 (1989).
15. H. Stoll, G. Wagenblast and H. Preuss, *Theor. Chim. Acta* **57**, 169-178 (1980).
16. M. Sironi, A. Genoni, M. Civera, S. Pieraccini and M. Ghitti, *Theor. Chem. Acc.* **117**, 685-698 (2007).
17. R. McWeeny, *Proc. R. Soc. London Ser. A* **253**, 242-259 (1959).
18. R. McWeeny, *Rev. Mod. Phys.* **32**, 335-369 (1960).
19. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press, London, UK (1992).
20. W. H. Adams, *J. Chem. Phys.* **34**, 89-102 (1961).
21. S. Huzinaga and A. A. Cantu, *J. Chem. Phys.* **55**, 5543-5549 (1971).
22. T. L. Gilbert, *J. Chem. Phys.* **60**, 3835-3844 (1974).
23. O. Matsuoka, *J. Chem. Phys.* **66**, 1245-1254 (1977).
24. G. G. Smits and C. Altona, *Theor. Chim. Acta* **67**, 461-475 (1985).
25. E. Francisco, A. M. Pendás, and W. H. Adams *J. Chem. Phys.* **97**, 6504-6508 (1992).
26. P. Ordejón, D. A. Drabold, M. P. Grumbach and R. M. Martin, *Phys. Rev. B* **48**, 14646-14649 (1993).
27. M. Couty, C. A. Bayse, and M. B. Hall, *Theor. Chem. Acc.* **97**, 96-109 (1997).
28. A. Fornili, M. Sironi and M.Raimondi, (2003). *J. Mol. Struct. (Theochem)* **632**, 157-172 (2003).
29. Z. Szekeres and P. R. Surján, *Chem. Phys. Lett.* **369**, 125-130 (2003).
30. D. M. Philipp and R. A. Friesner, *J. Comput. Chem.* **20**, 1468-1494 (1999).
31. B. Meyer, B. Guillot, M. F. Ruiz-López and A. Genoni, *J. Chem. Theory Comput.* **12,** 1052-1067 (2016).
32. B. Meyer, B. Guillot, M. F. Ruiz-López, C. Jelsch and A. Genoni, *J. Chem. Theory Comput.* **12**, 1068-1081 (2016).
33. D. Jayatilaka, *Phys. Rev. Lett.* **80**, 798–801 (1998).
34. D. Jayatilaka and D. J. Grimwood, *Acta Cryst. A* **57**, 76–86 (2001).
35. D. J. Grimwood and D. Jayatilaka, *Acta Cryst. A* **57**, 87–100 (2001).
36. I. Bytheway, D. J. Grimwood and D. Jayatilaka, *Acta Cryst. A* **58**, 232–243 (2002).

37. I. Bytheway, D. J. Grimwood, B. N. Figgis, G. S. Chandler and D. Jayatilaka, *Acta Cryst. A* **58**, 244-251 (2002).
38. D. J. Grimwood, I. Bytheway and D. Jayatilaka, *J. Comput. Chem.* **24**, 470–483 (2003).
39. D. Jayatilaka, in *Modern Charge-Density Analysis*, Eds. C. Gatti and P. Macchi, Springer, Berlin, **6**, 213-257 (2012).
40. M. Hudák, D. Jayatilaka, L. Perašínová, S. Biskupic, J. Kozísek and L. Bučinský, *Acta Cryst. A* **66**, 78–92 (2010).
41. B. Meyer and A. Genoni, *in preparation*.
42. D. Jayatilaka and B. Dittrich, *Acta Cryst. A* **64**, 383- 393 (2008).
43. S. C. Capelli, H.-B. Bürgi, B. Dittrich, S. Grabowsky and D. Jayatilaka, *IUCrJ*, **1**, 361-379 (2014).
44. M. Woińska, D. Jayatilaka, M. A. Spackman, A. J. Edwards, P. M. Dominiak, K. Woźniak, E. Nishibori, K. Sugimoto and S. Grabowsky, *Acta Cryst. A* **70**, 483-498 (2014).
45. M. Woińska, S. Grabowsky, P. M. Dominiak, K. Woźniak and D. Jayatilaka, *Sci. Adv.* **2**, e1600192 (2016).
46. A. Genoni, *J. Phys. Chem. Lett.* **4**, 1093-1099 (2013).
47. A. Genoni, *J. Chem. Theory Comput.* **9**, 3004-3019 (2013).
48. L. H. R. Dos Santos, A. Genoni and P. Macchi, *Acta Cryst. A* **70**, 532-551 (2014).
49. A. Genoni and B. Meyer, *Adv. Quantum. Chem.* **73**, 333-362 (2016).
50. A. Genoni, *Acta Cryst. A* **73**, 312-316 (2017).
51. N. Casati, A. Genoni, B. Meyer, A. Krawczuk and P. Macchi, *Acta Cryst. B* **73**, 584-597 (2017).
52. B. H. Chirgwin and C. A. Coulson, *Proc. R. Soc. London Ser. A* **201**, 196-209 (1950).
53. N. Casati, A. Kleppe, A. Jephcoat and P. Macchi, *Nat. Commun.* **7**, 10901 (2016).
54. L. Andrade Malaspina, S. Grabowsky and A. Genoni, *in preparation*.

Hirshfeld Atom Refinement

Simon Grabowsky
University of Bremen, Department 2 - Biology/Chemistry, Institute of Inorganic Chemistry and Crystallography, Bremen, Germany
simon.grabowsky@uni-bremen.de

Introduction

Hirshfeld atom refinement (HAR) [1, 2] is an advanced approach to X-ray crystal structure refinement. It is a post-IAM (IAM = Independent Atom Model) refinement procedure. HAR is available with the program Tonto (and its terminal version HARt, interfaced to Olex2).

In HAR, the atomic electron densities (ED) used to model the crystal ED during the refinement procedure are NOT spherical as in the IAM. Instead, they are deformed to account for aspherical effects, e.g. caused by bonding. They are calculated from the best reasonably achievable *ab-initio* quantum mechanical calculations (which relies on a choice of method and basis set).

The steps involved in HAR are:

1 - A theoretical single-point calculation provides the molecular electron-density distribution.

2 - This electron density undergoes

(i) a stockholder partitioning according to Hirshfeld [3, 4] to obtain the aspherical atomic electron densities which will then be

(ii) thermally smeared (by the convolution with the probability density function) and

(iii) Fourier transformed to provide tailor-made aspherical atomic scattering factors for the molecule being studied.

(i) $\rho(A) = w_A(\vec{r}) \cdot \rho_{molecule}(\vec{r})$,

Aspherical density of an atom at its position $\vec{r_A}$ in a molecule.

$$w_A(r) = \frac{\rho_A^0(\vec{r} - \vec{r_A})}{\sum_B \rho_B^0(\vec{r} - \vec{r_B})}$$

Hirshfeld or stockholder weight function where $\rho_B^0(\vec{r})$ is the spherically averaged atomic density for the isolated atom B centered at the origin.

(ii) $\langle \rho(A) \rangle = \rho_A * P_A$

Thermally smeared aspherical atomic electron density.

(iii) $F_j = \sum_a^N \sum_{\vec{r}} n_A^{-1} \int \langle \rho_a \rangle (\vec{r'}) e^{i\vec{q_j}.\vec{r'}} d\vec{r'}$

Aspherical structure factor, where N are the unique atoms that constitute an asymmetric unit. The factor n_A is a site-symmetry factor, i.e. the number of times the unique atom A is mapped onto itself.

3 - A full-matrix least-squares refinement of the parameters is then performed using the aspherical scattering factors in order to reproduce the observed X-ray diffraction data.

These steps are repeated until full convergence is achieved (in energy and geometry).
Since the wavefunction calculated in the energy computation in Tonto is based on an isolated molecule, the influence of the environment (or crystal effects) would not be reflected in the molecular electron density. Therefore, self-consistent atomic charges and dipoles of the Hirshfeld atoms are calculated and used to surround the isolated molecule for which the wavefunction is calculated to simulate a crystalline environment. The simulation of such effects by point charges and dipoles might be insufficient, e.g. for situations with strong hydrogen bonding, for such cases a more exhaustive treatment which explicitly includes the molecules involved in the strong interactions should be used.
Charges are placed on all complete molecules which have at least one atom within a specified distance of the central molecule. Complete molecules should be used, to avoid charged species and therefore minimize errors in slowly converging potential sums, but the option to use non-complete molecules is now available[1] and it is useful when dealing with network compounds.
It has been shown that including a surrounding cluster of charges is more important than using a large basis set,[2] and results for hydrogen atoms that are comparable to those obtained from neutron diffraction experiments might be achieved.[5, 6] When using HF/def2-SVP not including a self-consistent field of cluster charges surrounding the main molecule, this setting is called the "minimal HAR" [5] and is a good starting point.

Important facts:

1 - In the current version of Tonto, all hkl files must be merged and pruned of systematic absences. If you use HARt inside Olex2, this step is executed automatically by Olex2 with the use of cctbx.
2 - Since HAR is subjected to a theoretical single-point calculation for extraction of aspherical scattering factors, for structures with Z' < 1 the input geometry in the CIF must be modified in order to include one "full molecule" and/or "strong interactions" that might affect the ED.
3 - During the SCF computation in Tonto, all linear dependencies in the least-squares matrix are eliminated automatically, therefore no restrains or constraints are implemented in the program. This might cause problems in the case of spherical ions. To overcome this problem, we have created a script (lamaGOET – see below) that is able to perform a HAR (iteratively as the original HAR within Tonto) using different softwares for the SCF calculation (step 1), such as Gaussian, Orca[2], the original Tonto or the Elmo data base (by Alessandro Genoni). The lamaGOET script also provides a graphical interface which facilitates the generation of Tonto input files including advanced options.

1 This option was added by us in the Tonto code in Dec 4, 2017 (commit fe66cc2), and therefore does not work with earlier versions of Tonto.

2 The implementation with Orca is still not fully finalized since the current version of Tonto still cannot handle spherical basis-sets.

Instruction for HAR Workshop:

HARt in Olex2:

To perform HAR three files are needed:

1. CIF, containing the geometry of the structure to be refined
2. hkl file, containing all reflections measured
3. A basis set file, containing the coefficients for all atoms in your structure (default basis sets are provided with the program)

The most convenient way to get a HAR of your structure is through the Olex2 interface. For this purpose, you will need to install Olex2 from the website http://www.olexsys.org/. An .ins/.res/.cif-file can easily be loaded into Olex2 and the dropdown for HAR can be used to configure your HAR:

A structure can be either loaded by drag-and-drop of the corresponding structure file into the Olex2 window or using the File->Open dropdown menu. Now load the provided example of epoxide in the folder *input/epoxide_olex* or choose it in Olex2.

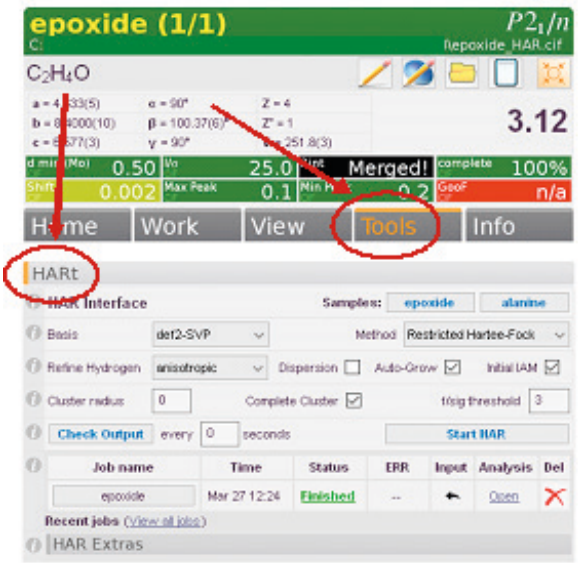


Figure 1.

When the structure is loaded and your IAM model is finished and converged (Shift turns green), you can go to Tools and open the HARt dropdown menu. A set of default options will be preconfigured. The recommendation for "Minimal HAR" is a level of theory of HF/def2-SVP without cluster charges or any other advanced setting, but anisotropic hydrogen treatment. Hit „Start HAR" and a window will open, where the output of the calculation is printed. This window must not be closed since it is linked with the calculation. When the calculation is finished you will see a massage like this:

```
Unit cell residual density:

Maximum ..................... 0.082212
- nearest atom ............... O1
- distance to it ............. 1.926179
Minimum ..................... -0.170281
- nearest atom ............... H3B
- distance to it ............. 0.744462
RMS ......................... 0.025979

Residuals on atoms:

Maximum ..................... 0.035162
- for atom ................... H2A
Minimum ..................... -0.051326
- for atom ................... H3B
RMS ......................... 0.024984

CPU time taken for job "epoxide" is 52.500 CPU seconds.
Wall-clock time taken for job "epoxide" is , 53 seconds, 89 milliseconds.

Finished
Drücken Sie eine beliebige Taste . . .
```

Figure 2.

If you now press a button, this window, if activated, will close. A line will be added to your job list, which looks like this:

Figure 3.

While a job is running, you can watch its progress if you press „Check Output". If you press „Check Output" after the first cycle of refinement, the name of the job will turn into a button. If you click on it, the preliminary results will be loaded. Keep in mind that these are just preliminary, the final geometry might still be different. When the calculation finishes, the line will change to:

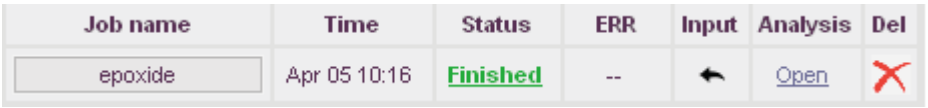


Figure 4.

If you click on „Finished" the output file will be opened, and you can browse through it. The button „Epoxide" opens the final structure. If you want to go back to the model before HAR the ↩ button will take you back.

If you have *plotly* installed for python in Olex2 (needs to be installed separately) you can click on the analysis „Open" link and it will open a tab in your browser showing interactive fitting-quality

plots. If you don't have it installed, the text file containing this information will be shown. If you want to remove a folder from your hard disk, hit „X". If you want to browse through your files, you can use the link „(View all jobs)" to open the folder containing all files from HAR runs. If you finished a calculation, the results will also be copied back into the folder where you started it. If you started it e.g. in D:\epoxide\ with a file epoxide.cif, the results will be called epoxide_HAR.cif. There is the .cif file containing the geometry, the .out file containing all output during the calculation and the .fcf and .fco files.

In cases where Z' < 1, Olex2 will try to grow your structure. In the examples folder *input/nh3_olex*, you will find a dataset for NH_3 which has only three atoms partially in the asymmetric unit. Using the tickbox „Auto-Grow" you can select whether Olex2 will try to grow your structure automatically. This works in most of the cases for molecular compounds such as NH_3, rubrene etc. If you hit „Start HAR" the molecule will be completed, and the refinement started. After a few seconds you will see the „Finished" message and can close the window with any key. If needed, hit „Check Output" and load the resulting cif. If you want to see the full structure you can use the toolbox provided on the bottom part of the HARt dropdown or use the default commands in Olex2 like *"grow"* or *"pack"*.

If you have a structure you need grow in a special, e.g. non-molecular, way the growing mode in Olex2 can be useful. Open the oxalate example and have a look at the hydrogen atom. It is involved in a strong hydrogen bond to a neighboring oxalate anion. If the wavefunction calculated for HAR does not have the information about the neighboring atom, the bond will be described in a different way than the crystal structure shows. Therefore, it is necessary to include the neighboring anion. This can be easily done using the growing tools provided in the HAR Extras.

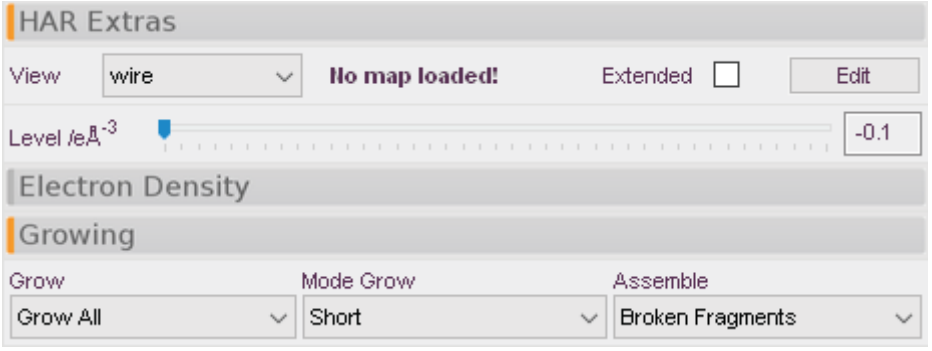


Figure 5.

Open them and select "Mode Grow" Short, as shown above. The structure will show all short contacts as dashed bonds. You can now click on the close contact of the hydrogen atom, and the second molecule will appear. Hit ESC to exit the growing mode.

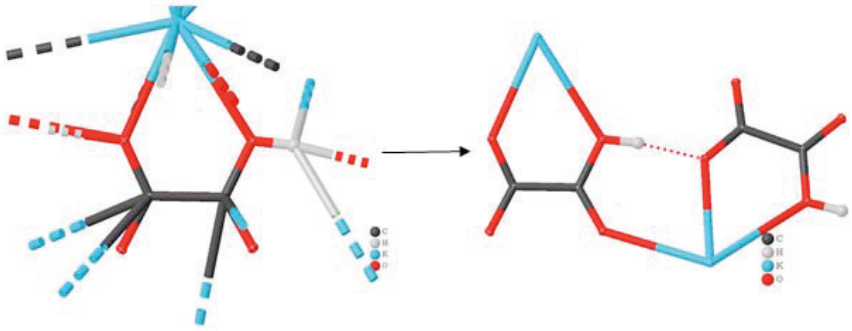


Figure 6.

If „Start HAR" is pressed now, the calculation will not start directly, a warning message is shown:

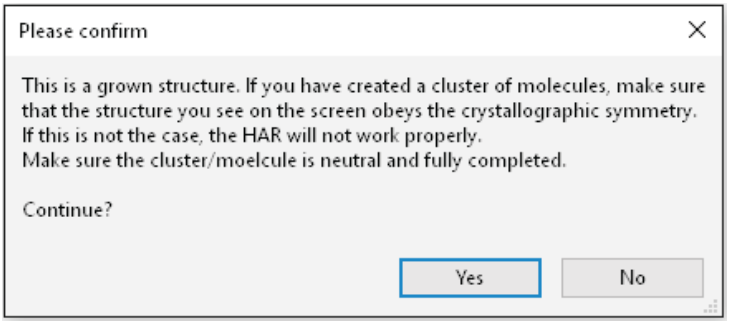


Figure 7.

If Yes is pressed, the calculation is started with both molecules for the wavefunction calculation, but only the asymmetric unit being refined in the least-squares. This will take longer than if only the asymmetric unit was considered, but the results are much more meaningful.

Manually running HARt:

If other features of HAR are needed, the executable hart(.exe) can be used. A full list of supported keywords can be obtained by typing:

$ ./hart.exe –help

A summary of information and a list of all possible keywords will be printed.
The default values and syntax are shown. The least amount of keywords essential for HARt can be derived by the necessary information for HAR as follows:

1. *-shelx-f / -shelx-f2 <FILENAME.hkl>*
2. *<FILENAME.cif>*

Everything else will be taken from default settings. The command needs to be executed in the folder containing the files or a full path will be needed for <FILENAME>. The output will be printed to <FILENAME.out> and basis sets will be taken from *./basis_sets*. If you want to

explicitly specify the location of a basis set directory the keyword *-basis-dir <path_to_basis_sets_folder>* will be needed, but in this case all example folders contain a basis_sets folder. Now open a terminal (on windows run cmd.exe or windows power shell, use windows+r and type cmd.exe or use the dropdown in the explorer) and navigate to the examples folder provided. Navigate to input/nh3_hart and type (depending on your Operating System):

> ..\hart.exe -shelx-f nh3.hkl nh3.cif *$ ../hart -shelx-f2 nh3.hkl nh3.cif*

The result files will be called nh3.archive.cif etc. You can open them using Olex2 or any other visualization tool you prefer. The result is identical to the one obtained by Olex2. However, if you want to have the most control over your calculation running HAR by calling *tonto* is the most sophisticated way.

Running HAR in Tonto:

Tonto is a program and library for quantum crystallography and quantum chemistry which allows you to:

- Calculate wavefunctions using Hartree-Fock, DFT, two-component relativistic methods, and perform analysis such as Roby bond analysis and property density plots.
- Refine crystal structures from structure factors using Hirshfeld atoms i.e. aspherical atomic scattering factors derived from ab initio wavefunctions.
- Obtain experimental wavefunctions that are constrained to reproduce experimental diffraction data.

Tonto is also used as the "back-end" to the popular CrystalExplorer visualization program. So performing HAR is only one of the features in Tonto.

The tonto executable is not provided together with Olex2 (unlike HARt) and most ideally should be installed and compiled on your operating system. For this workshop static release versions are provided.

To run HAR using the Tonto program itself, you will need to setup an extra file which is the input file for the Tonto job (stdin is the default name for this file). Using Tonto for the calculations allows access to several extra features compared to the terminal version HARt.

A good way to start with an input file is to copy one of the stdin files which already exist in the tests folder. Choose one which closely matches your type of job and modify it, but for this exercises an input file is provided in the *input/epoxide_tonto* folder.

The stdin input file always starts and ends with matching curly brackets, and it should contain instructions matching the hierarchy of Tonto.

The information regarding the format of your hkl file has to be added as well, this time with a header inside the hkl file instead of giving the format outside like in the HARt example.

Inside the folder *input/epoxide* provided, you will find the following files:

- epoxide.cif
- epoxide_with_header.hkl
- stdin

The *.cif-file* is identical to the ones used before in the previous examples, but the hkl file now includes a header to identify the columns in the hkl file.
The *stdin* contains all information needed for the refinement and all extra features of *tonto* are available through this file. The calculation is started with:

> .. \tonto.exe

A full list of all keywords available for each block in tonto can be printed in the stdout file by mistyping something inside the corresponding block in the stdin. There is a great number of keywords and inputs option for tonto that will be discussed but not explained explicitly here.

Running HAR in Tonto through the lamaGOET graphical interface, and extra options interfacing Tonto to different softwares:

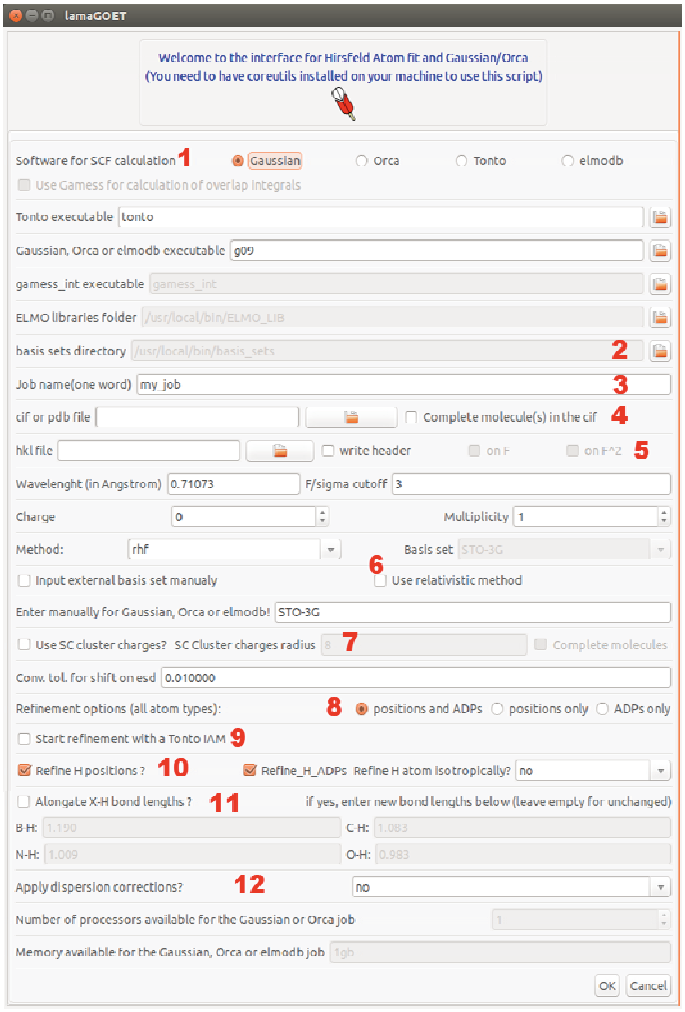


Figure 8.

Figure 1 shows a preview of the lamaGOET graphic interface where the numbers in red are:
1- Option to select different softwares for the single-point calculation.
2- Select the Tonto basis-set directory from the Tonto installation folder in your PC.
3- Job name which will be the new data block name in the resulting CIF and also in every file for the current refinement.
4- Option to be used in case of Z' < 1 to automatically complete the molecule in the asymmetric unit. Do not use if you have a network compound!
5- Option to automatically include the header required by Tonto in the reflection file (if you try to insert the header in a file that already contains it, the software will automatically ignore the request).
6- All basis sets available in Tonto are listed in a drop-down menu, but if you wish to use Gaussian for the energy calculation you need to enter the basis-set name manually in the Gaussian format.
7- Option to use self-consistent cluster charges during the calculation of the wavefunction. Point charges are calculated by Tonto from Hirshfeld partitioning, which is independent of the software chosen for the wavefunction calculation.
8- Refinement options.
9- Option to perform an IAM pre-refinement before HAR.
10- Option to refine or not refine the H positions. This option did not exist in Tonto until Dec 4, 2017.
11- Option to elongate X-H distances for all H atoms in the structure (independent of the hybridization state of the atom which it is bonded to).
12- Option to add dispersion corrections into the calculated structure factors. If set to yes, a popup window will be shown after pressing the OK button to enter f' and f" coefficients for each element that you wish to use the correction for. It is not required to enter these values for every element in the structure.
The script can be downloaded free of charge, for registered users, at https://sites.google.com/site/malaspinala/home/softwares.

1. S. C. Capelli, H.-B. Bürgi, B. Dittrich, S. Grabowsky, D. Jayatilaka: Hirshfeld atom refinement. *IUCrJ* **2014**, *1*, 361-379.
2. D. Jayatilaka, B. Dittrich, X-ray structure refinement using aspherical atomic density functions obtained from quantum-mechanical calculations. *Acta Cryst. A* **2008**, *64*, 383-393.
3. F. L. Hirshfeld, Can X-Ray Data Distinguish Bonding Effects from Vibrational Smearing. *Acta Cryst. A* **1976**, *32*, 239-244.
4. F. L. Hirshfeld, Bonded-Atom Fragments for Describing Molecular Charge-Densities. *Theor. Chim. Acta* **1977**, *44*, 129-138.
5. M. Fugel, D. Jayatilaka, E. Hupf, J. Overgaard, V. R. Hathwar, P. Macchi, M. J. Turner, J. A. K. Howard, O. V. Dolomanov, H. Puschmann, B. B. Iversen, H.-B. Bürgi, S. Grabowsky, Probing the accuracy and precison of Hirshfeld atom refinement with HARt interfaced with Olex2. *IUCrJ* **2018**, *5*, 32-44.
6. M. Woinska, S. Grabowsky, P. M. Dominiak, K. Wozniak, D. Jayatilaka, Hydrogen atoms can be located accurately and precisely by x-ray crystallography. *Sci. Adv.* **2016**, *2*, e1600192.

Atomic polarizabilities and dielectric properties

Piero Macchi
Department of Chemistry and Biochemistry University of Bern, Bern, Switzerland
piero.macchi@dcb.unibe.ch

The modern quantum chemical methods enable accurate calculations of molecular and crystal wavefunctions and, using response theory, of electronic properties.
However, for materials design, the periodic calculations are too expensive to allow screening structures that may offer the desired properties. For this reason, a simplification would be appreciated of a crystal in terms of building blocks, the assembling of which may return approximate properties, evaluated very rapidly. The building blocks may be molecules in a molecular crystal, isomeric unit of a polymer or even functional groups of any molecular base material (being polymeric or not). In fact, the concept of functional group is central in chemistry, adopted also to design molecules for, for example, drugs, pigments, chromophores, etc.
The optical, electric, magnetic or mechanical properties can be simplified in terms of a sum of functional group properties. This would enable for example, they consist of; ultimately, just one or a few functional groups might be enough to explain a molecular or crystal property, an important aspect for retro-designing new materials (*reverse crystal engineering*). The hypothesis is that functional groups retain their properties although embedded in different molecules (Hammet, 1937), at heart of the concept of *transferability* (Bader *et al.*, 1987; Bader *et al.* 1992). This approximation clearly neglects that the properties of a functional group depend also on the chemical environment, which perturbs the electron density of the functional group. We may distinguish through-bond and through-space perturbation, depending on conjugative effects or polarization induced by the electric field, respectively. The electric field due to the surrounding molecules in a crystal, despite being of the order of magnitude of GV/m, is much smaller than that of an intra-molecular field, thus, changing some substituents in a molecule is much more perturbative than changing the surrounding of a molecule.
In this lecture, the focus is on dielectric properties of materials that mainly depend on the electronic polarizability, in particular the first-order polarizability (*i.e.* the linear perturbation of the electron density distribution due to an electric field). The electronic polarizability is responsible for the *optoelectronic* properties of a material, like refractive index. A good rationalization of the key features that affect the polarizability allows one to design new materials with targeted applications, like high refractive index (HRI) materials are desirable for lenses, optical waveguides (Matsuda *et al.* 2000), high light extraction efficiency in LED components (Ju *et al.* 2006) lower dielectric constants for the insulation of wires (Hatton *et al.* 2006).
So far, the material design is based on empirically desumed building block properties. For example, aromatic rings, heteroatoms like sulphur, or halogen atoms were used to enhance

refractive indices. However, these *trial & error* attempts are time consuming, and do not allow the most efficient material design.

A rapid estimation of at least approximate values of dielectric constant of a molecular-based crystalline material is desirable to screen a much wider class of compound, without resorting on many measurements (not to mention the need itself of producing the material in adequate amount and form, which may be itself challenging).

An additional problem is calculating the materials properties with a sufficiently accurate theoretical method, going certainly beyond Hartree-Fock approximations (Champagne & Bishop, 2033). High correlation level is necessary to correctly represent local and non-local effects perturbing a functional group. However, correlated calculations with periodic boundary conditions are impossible. A possible solution is the partitioning that enables calculation at high level on a small fragment, which may recover the through-bond effects, whereas the through space interactions are simulated with semi-classical approaches. For this reason, a partition of the material in functional groups may be preferable with respect to computation of the polarizability of the entire system.

Thus, a database of functional groups may simultaneously guarantee the requested accuracy and low computational costs. A database of polarizabilities, and a protocol for the calculation of the atomic polarizabilities has been recently proposed, see Ernst *et al.* (2018).

Theoretical Background

As it is well known, a molecule in an electric field polarizes its charge density (in terms of both electron distribution and nuclear configuration). The (hyper)polarizabilities correlate the applied electric field and the perturbed electron density (Griffiths, 2008). In particular, the dipolar polarization (*i.e.* the induced change of dipole moment) is linearly dependent on the field, mediated by the first polarizability $\boldsymbol{\alpha}$:

$$\boldsymbol{\mu}_{induced} = \boldsymbol{\alpha}\mathbf{E} \tag{1}$$

where $\boldsymbol{\alpha}$ is a tensor and $\mathbf{E}$ is the vector field. Measurements, like the refraction in solution, enable determination of the isotropic value of the molecular polarizability), whereas more complicated is the measurement of individual components of the tensor.

In fact, for a linear, homogenous and linearly respondent dielectric, the Clausius-Mossotti equation holds true:

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{n^2 - 1}{n^2 + 2} = \frac{4\pi N\alpha}{3V} \tag{2}$$

where ε_r is the relative dielectric constant, n is the refractive index, and N is the number of molecules in the volume V. From equation (2), Pauling (1927) proposed the so-called *molar refraction* R

$$R = V\frac{n^2-1}{n^2+2} = \frac{4\pi N\alpha}{3} \qquad (3)$$

adopted even to determine the screening constants for atomic orbitals.

Many semi-empirical databases of atomic or group polarizabilities have been proposed, based on assumptions and using measurements like the refractive index.

Vogel (1948) reported on a database of isotropic polarizabilities for isolated atoms and functional groups using additivity schemes. Based on this scheme, Applequist (1976) proposed a non-additive model taking into account the induced dipole moments, thus introducing an anisotropy. Miller (1990) used the so-called atomic hybrid components, which takes into account the local hybridization of an atom, obtained by fitting to experimental data. Kassimi & Takkar (2009) introduced the concept of fragments in the calculation of organic molecules.

Stout & Dykstra (1995) developed a database of anisotropic polarizabilities obtained by fitting calculated molecular components.

One of the most important schemes was introduced by Stone (1985) who derived distributed atomic polarizabilities from distributed multipole expansion of the molecular charge density (Stone, 1981; Stone & Alderton, 1985; Le Sueur & Stone, 1994).

An important question is whether other experimental methods than molar refraction from solution can provide an experimental estimation of the molecular polarizabilities, in particular whether X-ray scattering. Ivanov-Smolenskii, Tsirelson & Ozerov (1983) proposed to compute the polarizabilities of ions from X-ray diffraction, but the most comprehensive work linking X-ray diffraction and molecular polarizability was that by Whitten, Spackman & Jayatilaka (2006). They pointed out that the one-electron density is not sufficient to calculate molecular polarizabilities, whereas the molecular orbitals from X-ray constrained wavefunction, with appropriate approximation, enable the estimation of molecular polarizabilities.

Another important question for the discussion in this lecture is the additivity of atomic / group polarizabilities (or electrostatic properties in general). Bader *et al.* (1992) demonstrated the additivity of group polarizabilities for acyclic hydrocarbons and rationalized it based on the partitioning of the electron density with the Quantum Theory of Atoms in Molecules (QTAIM). The scheme was further developed by Keith (2007) with a distribution of atomic dipole moments that solves (or better tackles) the origin dependence problem of dipole moment of non-neutral subunits of a molecule (like in fact atoms or functional groups). Keith's approach was further improved in the software *PolaBer*, which allows the calculation of full polarizability tensors for atoms and functional groups in isolated molecules or crystals (Krawczuk, Perez & Macchi, 2014)

and Dos Santos, Krawczuk & Macchi (2015) tested it on a series of amino acids in isolation or in aggregation. Moreover, this work paved the way for the compilation of a functional group database (Ernst *et al.*, 2018).

Distributed atomic polarizabilities

The molecular or crystal polarizability can be calculated by double derivative of the electronic energy with respect to the field, or by the first derivative of the dipole moment.

$$\alpha_{ij} = \left(\frac{\partial^2 E}{\partial \varepsilon_i \partial \varepsilon_j} \right) = \left(\frac{\partial \mu_i}{\partial \varepsilon_j} \right) \tag{4}$$

where E is the molecular energy and ε_i is a component of the electric field. Because of the symmetry of the second derivative, the polarizability tensor must be symmetric (thus $\alpha_{ij} = \alpha_{ji}$). Although the analytical differentiation is possible, often calculations make use of the finite field approach and the dipole moment is differentiated numerically.

For a calculation of atomic contributions, one can use a hard space partition (like Bader's) a fuzzy space partition (like for example Hirshfeld's) or a Hilber space partition. The hard space partition of the QTAIM is problematic because itself field dependent even for fixed nuclear geometries.

By adopting a finite field approach, the dipole moments of an atomic basins Ω are calculated and differentiated. It is interesting that the atomic dipole moments $\boldsymbol{\mu}(\Omega)$ consists of two terms: the atomic polarization $\mu_P(\Omega)$ (*i.e.*, the charge shift within the atomic basin) and the charge translation $\mu_C(\Omega)$ (*i.e.* the shift from one basin into another).

$$\mu(\Omega) = \mu_P(\Omega) + \mu_C(\Omega) = -\int_{\Omega} \left(\mathbf{r} - \mathbf{R}_{\Omega} \right) \rho(\mathbf{r}) d\mathbf{r} + \left(\mathbf{R}_{\Omega} - \mathbf{R}_0 \right) q(\Omega) \tag{5}$$

In equation (5), $\mathbf{R}_{\Omega}$ is the position vector of the nucleus Ω ; $\mathbf{R}_0$ is the origin of the molecular coordinate system (arbitrarily chosen). If the atomic basin is not neutral, $\mu_C(\Omega)$ does not vanish and is origin dependent. This may be a serious issue for transferability of an electrostatic property of an atom or a functional group in a molecule, given that they may not be neutral. Keith (2007) overcomes this problem transforming the charge-translation component into a basin-to-basin summation:

$$\mu(\Omega) = \mu_P(\Omega) + \mu_C(\Omega) = -\int_{\Omega} \left(\mathbf{r} - \mathbf{R}_{\Omega} \right) \rho(\mathbf{r}) d\mathbf{r} + \sum_{n_{BCP}} \left(\mathbf{R}_{\Omega} - \mathbf{R}_{BCP}(\Omega \,|\, \Omega') \right) q(\Omega \,|\, \Omega')$$

$q(\Omega \mid \Omega')$ is the charge transferred to the atomic basin Ω from the bonded atom Ω' and $\mathbf{R}_{BCP}$ is the positional vector of the bond critical point between atoms Ω and Ω'. Noteworthy, equation (5) is valid whatsoever definition of bond one adopts, not necessarily that of the QTAIM.
It follows that the atomic polarizabilities can by calculated as:

$$\alpha_{ij}(\Omega) = \lim_{\varepsilon_i \to 0} \frac{\mu_j^{\varepsilon_i}(\Omega) - \mu_j^{-\varepsilon_i}(\Omega)}{2\varepsilon_i} \tag{7}$$

If the space partition is an exact partition then the molecular polarizability is exactly reconstructed. However, because of the possible asymmetry of the atomic volume, the atomic polarizability is not symmetric, although the asymmetric components cancel each other when summing over all atoms. For this reason, a symmetrization of the atomic polaritabilities, according to the classical scheme by Nye (1985), one can remove the asymmetry of each atomic component, without affecting the total polarizability. Moreover, while the total molecular polarizability must be a positive definite tensor, the atomic polarizabilities must not. However, a negative atomic tensor may incidentally occur only for hydrogen atoms, especially when involved in strong hydrogen bonds and associated with a very low electronic population.
From equation (7) it immediately follows the possibility of computing functional group polarizabilities by adding atomic terms, as it was propose by Ersnt *et al.* (2018).

Dielectric properties from distributed atomic polarizabilities

One of the main applications of distributed atomic polarizabilities is the estimation of crystal optical properties, like dielectric constant:

$$\boldsymbol{\varepsilon}_{\mathbf{r}} = \boldsymbol{\chi} + 1 = \frac{4\pi \boldsymbol{\alpha}_{crystal}}{V} + 1 \tag{8}$$

There are different approaches, once a molecular polarizability is known (either from a direct *ab initio* calculation or from a database reconstruction).
A very crude approximation if the crystal polarizability is the simple sum of unperturbed molecular polarizabilities (which could be further decomposed in the sum of unperturbed atomic polarizabilities):

$$\boldsymbol{\alpha}^{crystal} = \sum_n \boldsymbol{\alpha}_n^{molecular} = \sum_n \sum_k \boldsymbol{\alpha}_{n.k}^{atomic} \tag{9}$$

Equation (9), however, does not take into account the enhancement of the polarizability in the field generated by other molecules. Within the classical electrostatics, this enhancement can be estimated by including the local field generated at a molecular / atomic site by the permanent

and induced dipole moments of the other molecules in the crystals (or atoms). Assuming the crystal an *ideal*, hence infinite, crystal, this calculation requires the Lorentz tensor **L**, following the suggestion by Dunmur (1972) and Cummins, Dunmur & Munn (1976):

$$\mathbf{E}(k) = \frac{1}{V}\sum_{k'}\mathbf{L}(k,k')\left(\boldsymbol{\mu}_0 + \boldsymbol{\mu}_{induced}\right) \tag{10}$$

where k, k' are lattice points of the ideal lattice. While the original definition was meaningful for an ideally homogeneous molecule occupying a lattice point of the crystal, Bounds & Munn (1977) have proposed a molecule as a collection itself of different sites (for example 6 sites in a benzene molecules), which turned out to be more accurate. Indeed, as shown by Krawczuk, Perez & Macchi (2014), this is suitable also for the case of distributed atomic polarizabilities. Bounds & Munn (1977) suggested to "distribute" the molecular polarizability of the central molecule on different sites, thus enabling a more accurate description. However, the distribution proposed was a simple equi-partition of the molecular tensor on some atomic sites, whereas it is clear from our examples that the atomic polarizabilities are quite diverse and this approach, albeit more precise than a central polarizability approach, may not be sufficiently accurate. Having, instead, the exact atomic polarizabilities, it is possible to perform the calculation with much more precision, using the same formalism of Bounds & Munn (1977).

The possibility of extracting the distributed polarizabilities, though, enables also additional treatments like for example the calculation of a molecule in the first coordination sphere, to account *ab initio* of the short intermolecular effects, and the subsequent calculation of Lorentz tensor, excluding the first coordination sphere. This provides an even better model, closer to the physical realm.

In all these approaches, anyway, a recursive calculation is necessary, given that the polarizabilities of the molecules at the lattice points is changing and therefore needs to be updated in a second cycle of calculations until convergence is reached.

References

Applequist, J. (1976) *Acc. Chem. Res.*, **10**, 79-85.

Bader, R. F. W., Keith, T. A., Gough, K. M. & Laidig, K. E. (1992). *Mol. Phys.*, **75**, 1167–1189.

Bader, R. F. W., Larouche, A., Gatti, C., Carroll, M. T., MacDougall, P. J. & Wiberg, K. B. (1987) *J. Chem. Phys.*, **87**, 1142-1152.

Bounds P. J. & Munn, R. W. (1977). *Chem. Phys.*, **24**, 343–353.

Champagne, B. & Bishop, D. M. (2003). *Adv. Chem. Phys.*, **126**, 41-92.

Dos Santos, L. H. R., Krawczuk, A. & Macchi, P. (2015). *J. Phys. Chem. A*, **119**, 3285–3298.

Cummins, P. G., Dunmur, D. A. / Munn, R. J. (1976). *Acta Cryst.*, **A32**, 847.

Dunmur, D. A. (1972). *Mol. Phys.*, **23**, 109.

Ju, Y., Almuneau, G., Kim, T. & Lee, B. (2006). *Jpn. J. Appl. Phys.*, **45**, 2546-2549.

Griffiths, D.J. (2008) *Introduction to Electrodynamics*, Pearson Benjamin Cummings, San Francisco, 3rd edition.

Hammett, L. P. (1937). *J. Am. Chem. Soc.*, **59**, 96-103.

Hatton, B. D., Landskron, K., Hunks, W. J., Bennett, M. R., Shukaris, D., Perovic, D. D. & Ozin, G. A. (2006). *Mater. Today*, **9**, 22–31.

Higashihara T. & Ueda, M. (2015). *Macromolecules*, **48**, 1915–1929.

Ivanov-Smolenskii, G. A., Tsirel'son, V. G. & Ozerov, R.P. (1983) *Acta Crystallogr. Sect. A*, **39**, 411-415.

Kassimi, N. E. B. & Thakkar, A. J. (2009) *Chem. Phys. Lett.*, **472**, 232–236.

Krawczuk, A., Pérez, D. & Macchi, P. (2014). *J. Appl. Crystallogr.*, **47**, 1452–1458.

Le Sueur, C. R. & Stone, A.J. (1994). *Mol. Phys.*, **83**, 293-307.

Liu, J. & Ueda, M. (2009). *J. Mater. Chem.*, **19**, 8907.

Macdonald, E.K. & Shaver, M. P. (2015). *Polym. Int.*, **64**, 6-14.

Matsuda, T., Funae, Y., Yoshida, M., Yamamoto, T. & Takaya, T. (2000). *J. Appl Polym. Sci.*, **76**, 45-49.

Miller, K. J. (1990). *J. Am. Chem. Soc.*, **112**, 8533–8542.

Nye, J.F. (1985). *Physical Properties of Crystals: Their Representation by Tensors and Matrices*, Oxford University Press, Oxford.

Pauling, L. (1927) *Proc. R. Soc. Lond.*, **114**, 181-211.

Stone, A.J. (1981) *Chem. Phys. Lett.,* **83**, 233-239

Stone, A.J. (1985) *Mol. Phys.*, **56**, 1065-1082.

Stone, A.J. & Alderton, M. (1985). *Mol. Phys.*, **56**, 1047-1064.

Stout, J. M. & Dykstra, C. E. (1995). *J. Am. Chem. Soc.*, **117**, 5127–5132.

Vogel, A. (1948). *J. Chem. Soc.*, **70**,1833-1855.

Wang, J., Hutchins, M., Woo, K., Konish, T. & Idacavage, M. J. (2010) *RadTech Report*, Sept-Oct., 11-17.

Whitten, A. E., Jayatilaka, D. & Spackman, M. A. (2006). *J. Chem. Phys.*, **125**, 174505.

Quantum Topology

Julia Contreras
CNRS-Sorbonne Universités, Paris, France
contrera@lct.jussieu.fr

A pillar of the emerging field of quantum crystallography is represented by the group of methods aiming at analyzing the chemical information contained in experimental and theoretical static electron density distributions.

The interest of this area of research is to focus on a classical concept that is extremely useful to understand and predict solid state behavior. As defined by Linus Pauling, "Chemistry is the science of substances: their structure, their properties, and the reactions that change them into other substances" [1]. The first two aspects, structure and properties, are clearly associated to the arrangement of atoms in a molecule, i.e., the chemical bond. These bonds determine chemical reactivity —the third aspect— and their visualization allows chemists to understand how atoms or molecules bond at a most fundamental level. A mechanistic understanding of chemical and biological functions and the structures of solid materials depends on knowing the geometric structures and the nature of bonds. But, despite the fact that the chemical bond is a fundamental concept in chemistry, "what is a chemical bond?" still remains a critical question for the chemical community because of the lack of a unique definition and inadequate understanding of its physical nature.

Visualization of bonding interactions between atoms and molecules is a long-standing quest in theoretical and computational chemistry. In recent years, it has become possible for most chemists to calculate molecular structures and stabilities based on quantum chemistry approaches. The main interest lies in creating a tool that enables researchers to see the interactions, and also interpret their characters and properties. Successful numerical solution of the Schrödinger equation yields energies and properties of atoms and molecules, but not directly a clear physical explanation of chemical bonding.

Theoretical framework

Several approaches have been developed to reveal the microscopic electronic natures of solids. Quantum Chemical Topology (QCT) does so by analyzing local functions, *f,* which yield a chemical picture of the system:

$$f : \mathbf{R}^3 \text{(molecular space)} \rightarrow \mathbf{R} \text{ (chemical picture)}$$

Generally, two different approaches are taken to analyze these functions:

- Local: it corresponds to the typical approach where the shape of a function is analyzed by looking at its maxima, minima,etc. In 3D, this means analyzing the points where the gradient of the function becomes zero (i.e. the critical points). There are 3 types of

critical points in 3D: maxima, minima, first order saddle points (maxima in 2 directions, minimal in 1) and second order saddle points (maximal in 1 direction, minima in 2). Their position and the value of *f* at these points helps understand the chemistry of the system.

- Global: the system is divided following surfaces of zero gradient. Taking a 2D example like in the Figure 1, the picture can be divided into two regions associated to the mountains.

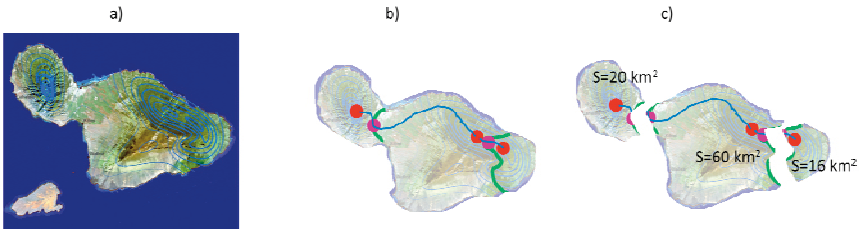


Figure 1. a) Island b) local information: Identification of the orography critical points c) global information: surface of each region induced by a topological partition.

As can be seen in Figure 1, this partition is exhaustive and non-overlapping. Hence, the sum of the different parts recovers the system (e.g. surface total of the island=96 km^2).
But how is this related to chemistry? The functions used need to have a chemical meaning. In what follows, we will see several examples of local functions which deliver different information on the system.

1. The electron density

The analysis of the topology of the electron density was introduced by Bader and collaborators in what is known as Quantum Theory of Atoms in Molecules (QTAIM) [2].
The electron density shows cusps (which can be associated with maxima). The first order saddle points (maxima in 2 directions, minimal in 1) are then associated with the presence of bonds in between pairs of atoms. This is why these critical points are also known as bond critical points (bcps) [3]. The position of the maxima (atoms) and bcps (bonds) thus enables to reconstruct the chemical graph of the molecule from calculated or experimental data. Figure 2 shows how this can be used to obtain the chemical graph of benzene. It has also enabled to solve controversial situations, such as the existence of a B-B bond in B_2H_6.
Since the partition we are using associated regions with the maxima, the final regions are associated with the nuclei. Hence, this partition enables to recover atomic regions. If we then integrate properties within these regions, we can obtain atomic regions. In this way, we can obtain atomic volumes and charges, which due to the properties of the partition, are additive and recover the system value. The transferability of these properties has enabled to recover the

transferability of functional groups.

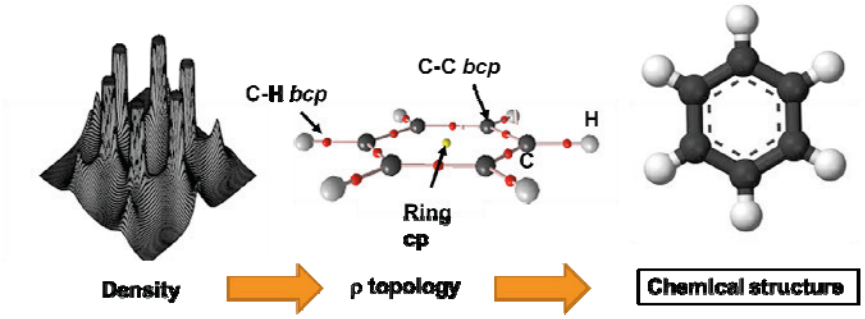


Figure 2. Scheme for obtaining the chemical graph of benzene from the electron density.

2. The Electron Localization Function

Another important set of *f* functions are those for the analysis of electron pairing, such as the electron density Laplacian,[4] the Electron Localization Function (ELF)[] and the Electron Localizability Indicator (ELI).[6]

The core of ELF, , can be understood as the Pauli kinetic energy density, t_P,scaled by the electron density, , so that it does not depend on the region of the molecule (the core has a much greater density).

$$\chi = \frac{t_P}{c_F \rho^{5/3}}$$

Then it is scaled to run from 0 to 1:

$$ELF = \frac{1}{1+\chi^2}$$

So that it shows maxima in the regions of electron pair localization. Hence, its maxima recover the Lewis structure: atomic shells, bonds and lone pairs. See Figure 3 with the ELF function for the water molecule: it shows the O-H bonds and the oxygen lone pairs.

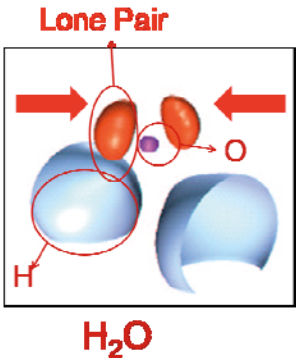


Figure 3. ELF picture of the water molecule.

The integration over these regions provides volumes for bonds and lone pairs. This analysis allows for example to rationalize the formation of channels in MOFs, clathrates or molecular crystals [7]. It is important to note that the ELF and ELI require the first-order density matrix, so that quantum crystallography developments also hold great potential for these analyses.

3. The reduced density gradient

Due to their delocalized nature, special *f* functions have been designed to visualize non-covalent interactions. As an example, the reduced density gradient (aka NCI for Non Covalent Interactions)[8], has been designed to detect weak interactions such as halogen bonds from the electron density. It helps to provide more stable pictures that do not change upon the quality of X-ray refinement. The reduced density gradient, s, depends on the electron density and its gradient:

$$s = \frac{|\nabla\rho|}{\rho^{4/3}}$$

It enables to visualize weak interactions, both attractive and repulsive. A continuous color-coding is used where strong and attractive non-covalent interactions, such as hydrogen bonds, are represented in blue, van der Waals interactions in green and repulsive interactions (steric clashes) in red (see Figure 4).

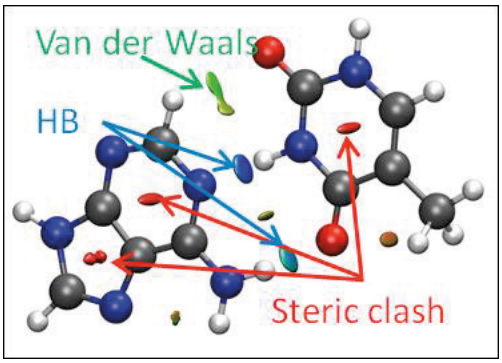


Figure 4. Non-covalent interactions in the adenine-thymine complex. Hydrogen Bonds in blue (2 intermolecular Hydrogen Bonds), steric clashes in red and dispersion in green.

A comparative view

In order to highlight the different chemical pictures provided by these functions, Figure xx shows the results for a benzene dimer extracted from the benzene crystal. QTAIM bcps reveal both covalent and intermolecular interactions in a benzene crystal. QTAIM, through integrations over atomic basins, yields atomic properties such as QTAIM charges (carbon and hydrogen charges, q, shown in Figure 5), atomic multipoles and volumes. The examination of BCPs provides insight into the structure and stability of crystals revealed in the bonding patterns that can be obtained from either theoretical or experimental electron densities. However, it is

not easy to differenciate the CH-C from the CH- interaction within this approach. The delocalized nature of CH- *vs* CH-C interactions becomes apparent when NCI is used. Finally, if we want to analyze the population of the bonding regions, and how they change upon the interaction, we need to analyze the ELF topology (Figure xx c).

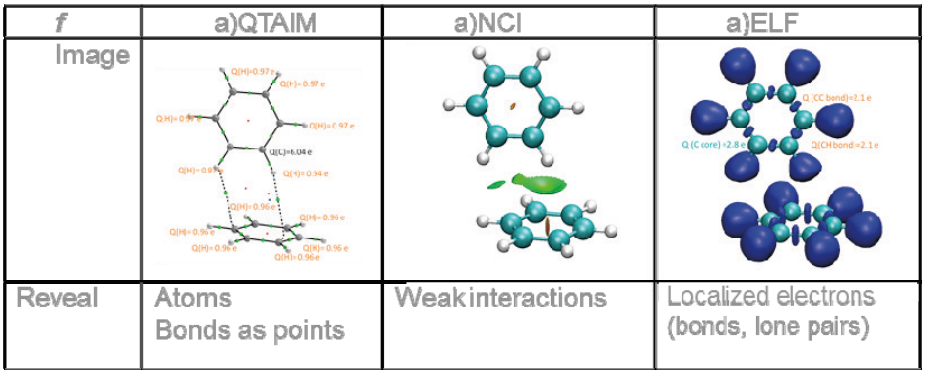


Figure 5. CH- and CH-C interactions in benzene crystal a) Atoms in Molecules, b) NCI, c) ELF

Applications

High pressure

Under pressure, properties of a material can get radically different from what they are at ambient pressure. For example, any material insulating at P = 1 atm is expected to turn metallic if submitted to a high enough pressure [9]. Maybe more appealing, unusual states of matter, such as superconductivity, can become quite common [10]. Concomitantly, the network of chemical bonds ensuring the cohesion of matter can get quite transformed as well as their very nature can change radically. As an example, weak intermolecular bonds that maintain together the molecules of a molecular solid can transform into strong "intramolecular" bonds under pressure, as the monomeric molecules that constitute the solid at low pressure polymerize into a 3D extended network of covalent bonds upon compression [11]. Typical examples are N_2 and CO_2 which are well known to form molecular solids at ambient pressure and transform into polymeric structures where N and C atoms are 3- and 4-coordinated at high pressure, respectively. These new coordinations correspond to those of P and Si in solid P and solid SiO_2 at atmospheric pressure. This is an illustration of the 9th Prewitt and Downs' rule of thumb which stipulates that elements behave at high pressures like the elements below them in the periodic table at lower pressures. Despite the establishment of many such rules, Hemley pointed out that "a fundamental yet empirically useful understanding of how pressure alters the chemistry of the elements is lacking". Therefore, pursuing the effort to build solid foundations for our chemical intuition under pressure is essential and the tools that we will see can help thereto.

We will cast the ability of the above tools to describe two types of phase transitions: related to crystal packing (physical), and related to bonding changes (chemical). Within the former, the capacity of pressure to promote
higher atomic coordinations in crystalline solids is one of its most outstanding features. This fact leads to densification processes of fundamental interest in areas ranging from planetary sciences to materials engineering. The electronic changes associated with a more effective atomic packing is an issue that needs to be addressed
if a complete characterization of the densification process is desired. In addition, the phase transition may be accompanied by a change in the general bonding pattern of the solid. In this case, understanding the process of bond formation and rupture becomes of crucial interest.

Property prediction

Topology has been a great source of insight in the understanding of crystalline organization. However it lacks a fundamental characteristic: its use have been barely predictive. This is so due to the lack of a direct (known) link between electron density topology and energetics. One way to approach this gap is to build energy models relying on topology. We have explored using a potential energy surface that includes chemical quantities explicitly, so that properties provided are directly related to the inherent organization of electrons within the regions provided by topological analysis. We will see how a very simple energetic model, the Bond Charge Model by Parr (BCM) [12], enables to describe the energetics of electron pairs. Coupling this to conceptual DFT, the band gap of solids can be univocally defined [13,14]. Applied to zinc-blende solids as a model case, trends in band gap can be predicted in terms of bond properties (length, charge, crystalline structure- Figure 6).

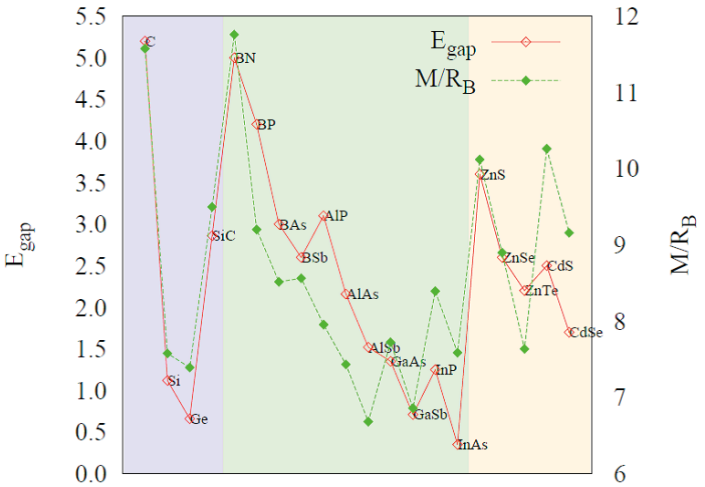


Figure 6. Band gap (in eV) for IV (purple), III-V (green) and I-VI (light orange) compounds from experimental data and from the ELF-BCM model.

References

[1] L. Pauling. *General Chemistry*. Dover Publications, New York (1947).

[2] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, Oxford, U.K., 1990.

[3] G. R. Runtz, R. F. W Bader, R. R. Messer, *Can. J. Chem*. **1977**, 55, 3040

[4] R. F. W. Bader, H. Essén, *J. Chem. Phys.* **1984**, *80*, 1943

[5] B. Silvi, A. Savin, *Nature* **1994**, *371*, 683

[6] M. Kohout, *Int. J. Quantum Chem.* **2004**, *97*, 651

[7] J. A. Sans, F. J. Manjon, A. L. Pereira, C. Popescu, A. Munoz, P. Rodriguez-Hernandez, J. Pellicer-Porres, V. P. Cuenca-Gotor, J. Contreras-García, V. Monteseguro-Padron, J. Ibanez, *in preparation*.

[8] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498-6506.

[9] E. Wigner, H. B. Huntington, *J. Chem. Phys.* **1935**, *3*, 764

[10] K. Amaya, K. Shimizu, M. I. Eremets, *Int. J. Mod. Phys. B* **1999**, *13*, 3623

[11] W. Grochala, R. Hoffmann, J. Feng, N. W. Ashcroft, *Angew. Chem. Int. Ed.* **2007**, *46*, 3620

[12] R. F. Borkman, R. G. Parr, *J. Chem. Phys*. **1968**, 48, 1116

[13] J. Contreras-Garcia, M. Marques, J.M. Menendez, J.M. Recio, *Int. J. Mol. Sci.* **2015**, 16, 8151

[14] J. Contreras-García, C. Cárdenas, *J Mol Mod* **2017**, 23, 271

Quantum refinement for biological applications

Kenneth M. Merz
Michigan State University, East Lansing, MI, USA
merzjrke@msu.edu

Ulf Ryde
Dept Theoretical Chemistry, Lund University, Lund, Sweden
ulf.ryde@teokem.lu.se

Quantum refinement is a method to supplement standard crystallographic refinement with quantum mechanical (QM) calculations.(Merz, 2015; Ryde, 2007; Ryde, Olsen, & Nilsson, 2002; Yu, Yennawar, & Merz, 2005) In the refinement process, the model (coordinates, *B* factors, occupancies, etc.) is optimized to provide an ideal fit to the experimental raw data (the structure factors).(Brünger & Rice, 1997; Kleywegt & Jones, 1997) The fit is measured by the crystallographic *R* factor or more sophisticated statistical measures ($E_{\text{X-ray}}$).(Adams, Pannu, Read, & Brünger, 1997) For the resolution obtained for most biological macromolecules, 1–3 Å, the available data is not enough to determine the exact position of all atoms. Therefore, the experimental data is supplemented by empirical chemical information in the form of ideal bond lengths, angles, dihedrals and non-bonded interactions.(Kleywegt & Jones, 1998) In the language of computational chemistry, this corresponds to a molecular-mechanics (MM) force field (E_{MM}),(Mackerell, 2004) although it is normally not energy-derived, but rather taken from a statistical survey of accurate crystal structures.(Engh & Huber, 1991) Consequently, the refinement takes the form of a minimization with the (pseudo-) energy function:(Jack & Levitt, 1978; Kleywegt & Jones, 1997)

$$E_{\text{cryst}} = E_{\text{X-ray}} + w_{\text{A}}\, E_{\text{MM}}$$

where $E_{\text{X-ray}}$ and E_{MM} are the crystallographic and MM energy functions and w_{A} is weight factor, needed because $E_{\text{X-ray}}$ and E_{MM} do not have the same units. The latter specifies the relative weight of the two terms and it is normally determined so that corresponding forces have a similar magnitude in a short molecular dynamics simulation of the crystal.(Brünger & Rice, 1997; Jack & Levitt, 1978)

This approach works fairly well for proteins and nucleic acids, for which there are plenty of information about the ideal geometry and a MM description works well.(Kleywegt & Jones, 1998) However, for other parts of the structure, e.g. metal sites, substrates, inhibitors, cofactors and ligands (which often are of prime interest for the function of the macromolecule), such information is normally missing or much less accurate.(Kleywegt & Jones, 1998; Mackerell, 2004;

Merz, 2015; Nilsson, Lecerof, Sigfridsson, & Ryde, 2003) Moreover, the MM description is rather inaccurate, typically missing, electrostatics, polarization and charge transfer.(Lopes et al., 2013; Mackerell, 2004) This can lead to serious errors in the final structure, often in the active sites of enzymes. This can be solved by employing a more accurate energy function, provided by QM calculations, which automatically involve all energy terms and does not require any parametrization.(Neese, 2006; Ryde, 2007) In the first implementation of this quantum-refinement approach, QM was only employed for a small, but interesting, part of the macromolecule, using the energy function(Ryde et al., 2002)

$$E_{\text{-X}} = E_{\text{X-ray12}} + w_{\text{A}} \left(w_{\text{QM}} E_{\text{QM1}} + E_{\text{MM12}} - E_{\text{MM1}}\right)$$

where E_{QM1} is the QM energy and the subscripts indicate if the method is used for the whole macromolecule ($_{12}$) or only for the QM regions ($_{1}$). In addition, another scale factor w_{QM} needs to be included, because the MM energy function is derived from statistics and therefore is typically 3 times larger than an energy-derived energy function.(Ryde et al., 2002).

This approach was implemented 2002 in the ComQum-X software (http://signe.teokem.lu.se/~ulf/Methods/comqum_x.html), using density-functional theory (DFT) calculations.(Ryde et al., 2002) It was shown to locally improve the geometry of metal sites in proteins,(Ryde & Nilsson, 2003) owing to the accurate structures provided by DFT.(Neese, 2006; Ryde, 2007) Moreover, it was shown that the protonation state of metal-bound ligands could be determined by comparing quantum-refined structures optimized in different protonation states with respect to the *R* factors, real-space *R* factors, bond lengths and the strain energy, i.e. the energy of the QM system when optimized in the protein and in vacuum (the latter giving the intrinsic geometry and energy).(Nilsson & Ryde, 2004) In the same way, the oxidation state of metal sites could be deduced,(Rulíšek & Ryde, 2006) but it was soon discovered that the oxidation state often changes during data collection,(Rulíšek & Ryde, 2006; Söderhjelm & Ryde, 2006) owing to photoreduction by electrons released by the X-rays.(H. P. Hersleth & Andersson, 2011).

For the first full protein refinement a related approach was used where E_{MM} in equation 1 was replaced by E_{QM} and the CNS refinement package was used.(Yu et al., 2005) A linear-scaling semiempirical QM method was used for the QM model in the re-refinement.(Dixon & Merz, 1997) This work involved the re-refinement of the BPTI structure (PDBID:5PTI) at 1Å resolution. Remarkably even at this high resolution the QM based re-refinement was able to correct several structural anomalies seen in the original classic refinement process.

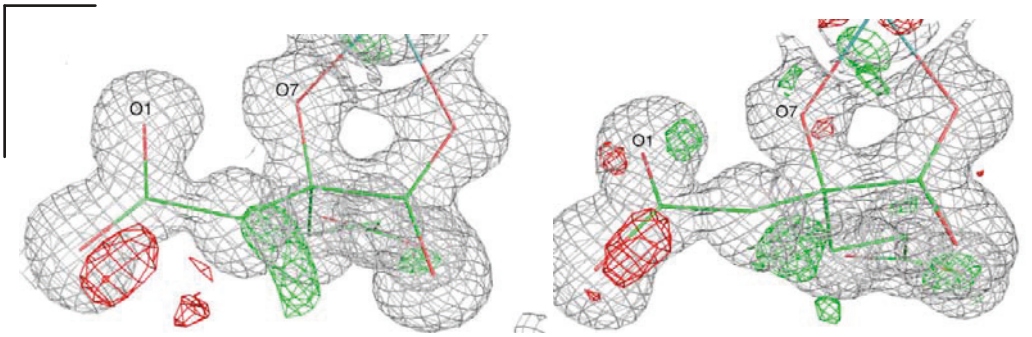


Figure 1. Electron-density maps of two possible protonation states of the homocitrate ligand in nitrogenase.(Cao, Caldararu, & Ryde, 2017) The $2mF_o - DF_c$ maps are contoured at 1.0 σ and the $mF_o - DF_c$ maps are contoured at +3.0 σ (green) and −3.0 σ (red). The figure to the left is better, especially around the O1 and O7 atoms, which is also reflected in the real-space difference density Z-scores,(Tickle, 2012) which are 3.0 and 3.2, respectively.

Quantum refinement has been applied to many systems of biological or chemical interest, e.g. ferrochelatase, cytochrome c_{553}, myoglobin, peroxidase, alcohol dehydrogenase, nitrite reductase, Mn superoxide dismutase, [NiFe] hydrogenase, nitrogenase, sulfite oxidase, particulate methane monooxygenase and zinc metalloproteases.(Caldararu, Andrejić, Cioloboc, & Ryde, 2017; Cao, Caldararu, Rosenzweig, & Ryde, 2017; Cao, Caldararu, & Ryde, 2017; Heimdal, Rydberg, & Ryde, 2008; H.-P. Hersleth, Hsiao, Ryde, Görbitz, & Andersson, 2008; Källrot, Nilsson, Rasmussen, & Ryde, 2005; Xue Li, Hayik, & Merz, 2010; Nilsson, Hersleth, Rod, Andersson, & Ryde, 2004; Nilsson & Ryde, 2004; Rulíšek & Ryde, 2006; Ryde & Nilsson, 2003; Ryde et al., 2002; Söderhjelm & Ryde, 2006).

Typical applications regard the nature, protonation and oxidation state of the active site, comparing different structural alternatives, as is shown in Figure 1. Application to drug design and ligand refinement(Borbulevych, Plumley, Martin, Merz, & Westerhoff, 2014; Fu, Li, & Merz, 2012; Fu, Li, Miao, & Merz, 2013; Xue Li, Fu, & Merz, 2011; Xue Li, He, Wang, & Merz, 2009; Yu, Hayik, et al., 2006; Yu, Li, Cui, Hayik, & Merz, 2006) have advanced the accuracy of the representation of small-molecules in active site pockets over what is possible with classical models where the force field representing the ligand is generally less well validated than that for the protein. In the latter studies the E_{MM} term in equation 1 was replaced with $E_{QM/MM}$ where the QM model was either semiempirical or *ab initio* and was implemented in the AMBER suite of programs. Naturally, the largest effects in most cases were seen for low resolutions and at resolutions better than ~1 Å, effects of systematic errors in the QM method start to be apparent. The method has been extended to neutron structures,(Manzoni, Caldararu, Oksanen, Logan, & Ryde, 2017) as well as to NMR structure refinement(Chakravorty et al., 2013; He, Wang, & Merz, 2009; Hsiao, Drakenberg, & Ryde, 2005; Wang & Merz Jr, 2005; Wang, Raha, & Merz, 2004; Wang, Westerhoff, & Merz, 2007; Williams, Peters, Wang, Roitberg, & Merz, 2009) and normal or polarized EXAFS (extended X-ray absorption fine structure) refinement.(Hsiao & Ryde, 2006; Hsiao, Tao, Shokes, Scott, & Ryde, 2006; Xichen Li, Siegbahn, & Ryde, 2015; Xichen Li,

Sproviero, Ryde, Batista, & Chen, 2013; Ryde, Hsiao, Rulíšek, & Solomon, 2007) In addition, the various methods can be combined.(Xichen Li et al., 2013; Manzoni et al., 2017)

References

Adams, P. D., Pannu, N. S., Read, R. J., & Brünger, A. T. (1997). Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(10), 5018–5023. http://doi.org/10.1073/pnas.94.10.5018

Borbulevych, O. Y., Plumley, J. A., Martin, R. I., Merz, K. M., & Westerhoff, L. M. (2014). Accurate macromolecular crystallographic refinement: Incorporation of the linear scaling, semiempirical quantum-mechanics program DivCon into the PHENIX refinement package. *Acta Crystallographica Section D: Biological Crystallography*, *70*(5), 1233–1247. http://doi.org/10.1107/S1399004714002260

Brünger, A. T., & Rice, L. M. (1997). Crystallographic Refinement by Simulated Annealing: Methods and Applications. *Methods in Enzymology*, *277*, 243–269.

Caldararu, O., Andrejić, M., Cioloboc, D., & Ryde, U. (2017). QM/MM study of sulfite oxidase. *Journal of Biological Inorganic Chemistry*, submitted.

Cao, L., Caldararu, O., Rosenzweig, A. C., & Ryde, U. (2017). Quantum refinement gives no support to dinuclear copper sites in crystal structures of particulate methane monooxygenase. *Angewandte Chemie (International Ed.)*, *submitted*.

Cao, L., Caldararu, O., & Ryde, U. (2017). Protonation states of homocitrate and nearby residues in nitrogenase studied by computational methods and quantum refinement. *Journal of Physical Chemistry B*, *in press*, doi: 10.1021/acs.jpcb.7b02714.

Chakravorty, D. K., Wang, B., Lee, C. W., Guerra, A. J., Giedroc, D. P., & Merz, K. M. (2013). Solution NMR refinement of a metal ion bound protein using metal ion inclusive restrained molecular dynamics methods. *Journal of Biomolecular NMR*. http://doi.org/10.1007/s10858-013-9729-7

Dixon, S. L., & Merz, K. M. (1997). Fast, accurate semiempirical molecular orbital calculations for macromolecules. *The Journal of Chemical Physics*, *107*(3), 879. http://doi.org/10.1063/1.474386

Engh, R. A., & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallographica Section A*, *47*(4), 392–400. http://doi.org/10.1107/S0108767391001071

Fu, Z., Li, X., & Merz, K. M. (2012). Conformational Analysis of Free and Bound Retinoic Acid. *Journal of Chemical Theory and Computation*, *8*(4), 1436–1448. http://doi.org/10.1021/ct200813q

Fu, Z., Li, X., Miao, Y., & Merz, K. M. (2013). Conformational analysis and parallel QM/MM X-ray refinement of protein bound anti-Alzheimer drug donepezil. *Journal of Chemical Theory and Computation*, *9*(3), 1686–1693. http://doi.org/10.1021/ct300957x

He, X., Wang, B., & Merz, K. M. (2009). Protein NMR chemical shift calculations based on the automated fragmentation QM/MM approach. *The Journal of Physical Chemistry. B*, *113*(30), 10380–8. http://doi.org/10.1021/jp901992p

Heimdal, J., Rydberg, P., & Ryde, U. (2008). Protonation of the proximal histidine ligand in heme peroxidases. *Journal of Physical Chemistry B*, *112*(8), 2501–2510. http://doi.org/10.1021/jp710038s

Hersleth, H.-P., Hsiao, Y.-W., Ryde, U., Görbitz, C. H., & Andersson, K. K. (2008). The crystal structure of peroxymyoglobin generated through cryoradiolytic reduction of myoglobin compound III during data collection. *Biochem. J.*, *412*(2), 257–64. http://doi.org/10.1042/BJ20070921

Hersleth, H. P., & Andersson, K. K. (2011). How different oxidation states of crystalline myoglobin are

influenced by X-rays. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1814*, 785–796. http://doi.org/10.1016/j.bbapap.2010.07.019

Hsiao, Y. W., Drakenberg, T., & Ryde, U. (2005). NMR structure determination of proteins supplemented by quantum chemical calculations: Detailed structure of the Ca2+ sites in the EGF34 fragment of protein S. *Journal of Biomolecular NMR*, *31*(2), 97–114. http://doi.org/10.1007/s10858-004-6729-7

Hsiao, Y. W., & Ryde, U. (2006). Interpretation of EXAFS spectra for sitting-atop complexes with the help of computational methods. *Inorganica Chimica Acta*, *359*(4), 1081–1092. http://doi.org/10.1016/j.ica.2005.11.036

Hsiao, Y. W., Tao, Y., Shokes, J. E., Scott, R. A., & Ryde, U. (2006). EXAFS structure refinement supplemented by computational chemistry. *Physical Review B - Condensed Matter and Materials Physics*, *74*(21), 1–17. http://doi.org/10.1103/PhysRevB.74.214101

Jack, A., & Levitt, M. (1978). Refinement of large structures by simultaneous minimization of energy and R factor. *Acta Crystallographica Section A*, *34*(6), 931–935. http://doi.org/10.1107/S0567739478001904

Källrot, N., Nilsson, K., Rasmussen, T., & Ryde, U. (2005). Theoretical study of structure of catalytic copper site in nitrite reductase. *International Journal of Quantum Chemistry*, *102*(5 SPEC. ISS.), 520–541. http://doi.org/10.1002/qua.20386

Kleywegt, G. J., & Jones, T. A. (1997). Model Building and Refinement Practice. *Methods in Enzymology*, *277*, 208–230.

Kleywegt, G. J., & Jones, T. A. (1998). Databases in protein crystallography. *Acta Crystallographica Section D: Biological Crystallography*, *54*(6 I), 1119–1131. http://doi.org/10.1107/S0907444998007100

Li, X., Fu, Z., & Merz, K. M. (2011). QM/MM refinement and analysis of protein bound retinoic acid. *Journal of Computational Chemistry*, 1–10. http://doi.org/10.1002/jcc.21978

Li, X., Hayik, S. a, & Merz, K. M. (2010). QM/MM X-ray refinement of zinc metalloenzymes. *Journal of Inorganic Biochemistry*, *104*(5), 512–22. http://doi.org/10.1016/j.jinorgbio.2009.12.022

Li, X., He, X., Wang, B., & Merz, K. (2009). Conformational variability of benzamidinium-based inhibitors. *Journal of the American Chemical Society*, *131*(22), 7742–54. http://doi.org/10.1021/ja9010833

Li, X., Siegbahn, P. E. M., & Ryde, U. (2015). Simulation of the isotropic EXAFS spectra for the S2 and S3 structures of the oxygen evolving complex in photosystem II. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 3979–84. http://doi.org/10.1073/pnas.1422058112

Li, X., Sproviero, E. M., Ryde, U., Batista, V. S., & Chen, G. (2013). Theoretical EXAFS studies of a model of the oxygen-evolving complex of photosystem II obtained with the quantum cluster approach. *International Journal of Quantum Chemistry*, *113*(4), 474–478. http://doi.org/10.1002/qua.24143

Lopes, P., Huang, J., Shim, J., Luo, Y., Li, H., Roux, B., & MacKerell, A. (2013). Polarizable Force Field of Peptide and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.*, *9*(12), 5430–5449.

Mackerell, A. D. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, *25*(13), 1584–1604. http://doi.org/10.1002/jcc.20082

Manzoni, F., Caldararu, O., Oksanen, E., Logan, D. T., & Ryde, U. (2017). Refinement of protein structures using a combination of quantum mechanical calculations with neutron and X-ray crystallographic data. *Journal of Applied Crystallography*, *submitted*.

Merz, K. M. (2015). Using Quantum Mechanical Approaches to Study Biological Systems.

Neese, F. (2006). A critical evaluation of DFT, including time-dependent DFT, applied to bioinorganic chemistry. *Journal of Biological Inorganic Chemistry*, *11*(6), 702–711. http://doi.org/10.1007/s00775-

006-0138-1
Nilsson, K., Hersleth, H.-P., Rod, T. H., Andersson, K. K., & Ryde, U. (2004). The protonation status of compound II in myoglobin, studied by a combination of experimental data and quantum chemical calculations: quantum refinement. *Biophysical Journal*, *87*(5), 3437–47. http://doi.org/10.1529/biophysj.104.041590
Nilsson, K., Lecerof, D., Sigfridsson, E., & Ryde, U. (2003). An automatic method to generate force-field parameters for hetero-compounds. *Acta Crystallographica - Section D Biological Crystallography*, *59*, 274–289. http://doi.org/10.1107/S0907444902021431
Nilsson, K., & Ryde, U. (2004). Protonation status of metal-bound ligands can be determined by quantum refinement. *Journal of Inorganic Biochemistry*, *98*(9), 1539–1546. http://doi.org/10.1016/j.jinorgbio.2004.06.006
Rulíšek, L., & Ryde, U. (2006). Structure of reduced and oxidized manganese superoxide dismutase: A combined computational and experimental approach. *Journal of Physical Chemistry B*, *110*(23), 11511–11518. http://doi.org/10.1021/jp057295t
Ryde, U. (2007). Accurate metal-site structures in proteins obtained by combining experimental data and quantum chemistry. *Dalton Transactions (Cambridge, England 2003)*, (6), 607–25. http://doi.org/10.1039/b614448a
Ryde, U., Hsiao, Y. W., Rulíšek, L., & Solomon, E. I. (2007). Identification of the peroxy adduct in multicopper oxidases by a combination of computational chemistry and extended X-ray absorption fine-structure measurements. *Journal of the American Chemical Society*, *129*(4), 726–727. http://doi.org/10.1021/ja062954g
Ryde, U., & Nilsson, K. (2003). Quantum Chemistry Can Locally Improve Protein Crystal Structures. *Journal of the American Chemical Society*, *125*(47), 14232–14233. http://doi.org/10.1021/ja0365328
Ryde, U., Olsen, L., & Nilsson, K. (2002). Quantum chemical geometry optimizations in proteins using crystallographic raw data. *Journal of Computational Chemistry*, *23*(11), 1058–1070. http://doi.org/10.1002/jcc.10093
Söderhjelm, P., & Ryde, U. (2006). Combined computational and crystallographic study of the oxidised states of [NiFe] hydrogenase. *Journal of Molecular Structure: THEOCHEM*, *770*(1–3), 199–219. http://doi.org/10.1016/j.theochem.2006.06.008
Tickle, I. J. (2012). Statistical quality indicators for electron-density maps. *Acta Crystallographica Section D Biological Crystallography*, *68*, 454–467. http://doi.org/10.1107/S0907444911035918
Wang, B., & Merz Jr, K. M. (2005). Validation of the binding site structure of the cellular retinol-binding protein (CRBP) by ligand NMR chemical shift perturbations. *J. Am. Chem. Soc*, *127*(15), 5310–5311. Retrieved from http://pubs.acs.org/doi/abs/10.1021/ja042616k
Wang, B., Raha, K., & Merz, K. M. (2004). Pose scoring by NMR. *Journal of the American Chemical Society*, *126*(37), 11430–1. http://doi.org/10.1021/ja047695e
Wang, B., Westerhoff, L. M., & Merz, K. M. (2007). A critical assessment of the performance of protein-ligand scoring functions based on NMR chemical shift perturbations. *Journal of Medicinal Chemistry*, *50*(21), 5128–34. http://doi.org/10.1021/jm070484a
Williams, D. E., Peters, M. B., Wang, B., Roitberg, A. E., & Merz, K. M. (2009). AM1 parameters for the prediction of 1H and 13C NMR chemical shifts in proteins. *The Journal of Physical Chemistry. A*, *113*(43), 11550–9. http://doi.org/10.1021/jp9028722
Yu, N., Hayik, S. a, Wang, B., Liao, N., Reynolds, C. H., & Merz, K. M. (2006). Assigning the protonation

states of the key aspartates in beta-Secretase using QM/MM X-ray structure refinement. *Journal of Chemical Theory and Computation*, *2*(4), 1057–1069. http://doi.org/10.1021/ct0600060

Yu, N., Li, X., Cui, G., Hayik, S. a, & Merz, K. M. (2006). Critical assessment of quantum mechanics based energy restraints in protein crystal structure refinement. *Protein Science: A Publication of the Protein Society*, *15*(12), 2773–84. http://doi.org/10.1110/ps.062343206

Yu, N., Yennawar, H. P., & Merz, K. M. (2005). Refinement of protein crystal structures using energy restraints derived from linear-scaling quantum mechanics. *Acta Crystallographica Section D: Biological Crystallography*, *61*(3), 322–332. http://doi.org/10.1107/S0907444904033669

Experimental charge density studies in biomolecules

Benoit Guillot
CRM2 Lab - Lorraine University, Faculté des Sciences et Technologies, Vandoeuvre-lès-Nancy, France
benoit.guillot@univ-lorraine.fr

Experimental charge density science and the study of properties of biomolecules are deeply related. A week of intensive course in this matter (this lecture being scheduled the penultimate day of this Erice session) hopefully convinced the reader that there is a lot to learn from the knowledge of the molecular charge density, whether it is known from theoretical methods, from a high-resolution diffraction experiment or from a combination of both such as in "Quantum Crystallography" approaches.

This appears especially true in the specific case of biomolecules. Indeed, by definition, biomolecules are either naturally found in living organisms (endogenous), or presenting a biological activity (exogenous, for instance pharmaceutical drugs). These biological activities are necessarily related to their chemical properties and thus, to their charge distributions. There is abundant literature on experimental charge density studies of small biomolecules, both endogenous and exogenous. Indeed, the physicochemical knowledge one can expect from such experiments range from atomic charges, molecular dipole moments and other electrostatic properties (such as electrostatic potential indicating nucleophilic and electrophilic sites of a molecule), to topological indicators as defined in the framework of the QTAIM theory. All these properties can be related to the biological activity of the studied molecule, including its chemical reactivity or its ability to bind or to interact with a biological partner such as a protein (enzymes, transport proteins etc..).

In my lecture however, I will focus on the application of experimental charge density methods to study properties of *macromolecular systems*. Hence, concerning studies of "small" biological molecules, I will limit myself to give the course participants a representative (but certainly not exhaustive) bibliography. These references [1-75] are given with their titles and classified by types of biomolecules, then in chronological order. It appears obviously that amino acids and oligopeptides were the first kind of biomolecules raising interests in the experimental charge density science community [1-28], with clear perspectives of possible applications to the study of protein structures. It must be noted however that studied amino acids are not necessarily in their natural L-forms: racemic DL mixtures allowing, of course, the formation of centrosymmetric crystals. Literature is a bit scarcer about nucleic acids (nucleobases, nucleosides or nucleotides), or nucleic acids derivatives [29-41], despite the quite large representation of crystal structures related to nucleic acids in the Cambridge Structural Database. Other types of endogenous biomolecules have seen their electron density experimentally measured [42-52]: hormones,

neurotransmitters, carbohydrates, even the very large vitamin B_{12} (more than 250 atoms in the asymmetric unit!) has been successfully studied, using a combination of transferred aspherical scattering factors and theoretical computations to characterize its cobalt center [49]. A number of pharmaceutical drugs have also been considered [53-75], with often the perspective of determining relationships between their electron density properties and their biological activities (for instance [60, 61]). Again, I wish to highlight the experimental charge density studies of unusually large molecules of biological interest, such as the immunosuppressant drug cyclosporine [65], a cyclic peptide of nearly 200 atoms, or the one of trichotoxin_A50E [67], an antibiotic peptide made of 18 non-proteinogenic amino acids. Finally, I suggest few reviews in the field of charge density science applied to biology, QSAR, drug design or medicinal chemistry [76-79].

I now briefly summarize the topics I will actually discuss in my lecture, and again give a representative associated bibliography. I will first describe the specificities of macromolecular bio-crystallography and of protein structures, then introduce methods and computer programs that have been developed to extend what is being done in "traditional" charge density analyses (of small molecules) to the field of structural biology. Then I will review some of the most significant experimental charge densities analyses of protein structures published so far. I will base my lecture on selected examples among the ones I mention in these notes, with the objective of giving the audience the current state of the art in this field.

Ultra-high resolution protein structures

Protein crystals are difficult to obtain and result from a multidisciplinary process involving biochemistry (cloning, expression, purification of the chosen protein), crystallization experiments, diffraction data collection, data phasing, structure solution, and finally model refinement methods that are specific to this field [80]. Resulting crystals are fragile, usually of poor quality in terms of diffracting power (at least from a small molecules crystallographer point of view) and are characterized by high levels (~ 50% in average) of disordered solvent content [81]. Another characteristic of protein structures is the almost systematic presence of disorder. Proteins are indeed very flexible molecules, and this flexibility shows up as alternate conformations of amino acids side chains, or even main chain especially at high resolution. This can also be seen, when X-ray data are collected at cryogenic temperature, as "frozen" static disorder leading to atomic equivalent *B* factors in protein structures significantly larger than the ones found in isolated amino acids crystals studied at higher temperatures. Hence, despite progresses in X-ray detection technologies, crystallogenesis methods (robots!) and the advent of new generation synchrotrons, only 34 structures of proteins or peptides larger than 10 residues have been deposited in the Protein Data Bank [82] at resolutions strictly better than 0.8Å. Among these ultra-high resolution protein structures, a handful only has been studied by the mean of charge density science approaches (Table 1).

PDB ID	Resolution (Å)	Number of residues	Number of atoms	Macromolecule Name
5D8V	0.48	83	920	High-potential iron-sulfur protein
3NIR	0.48	48	470	Crambin
1EJG	0.54	48	424	Crambin
4HP2	0.64	38	395	Thiostrepton peptide
2VB1	0.65	129	1375	Lysozyme
1US0	0.66	316	3066	Aldose Reductase
4REK	0.74	499	4290	Cholesterol oxidase
5GV8	0.78	272	2490	NADH-cytochrome b5 reductase 3

Table 1. protein structures deposited in the Protein Data Bank at X-ray diffraction data resolution better than 0.8Å which have been studied by the mean of charge density science approaches.

Methods in charge density studies: from small molecules to proteins

The transferability principle (initially exposed in a paper by Brock *et al.* in 1991 [83]), and the development of the associated libraries of transferable electron density fragments undoubtedly contributed to the early interest of the charge density community for protein structures [84-86]. The transferability principle is based on the fact that the electron density parameters of an atom (as defined in the Hansen & Coppens multipole model [87]) in a particular environment are nearly the same in all compounds containing this atom. Based on this observation, three libraries of electron density "building blocks" have been developed, either from accurate experimental charge density studies of small molecules (*ELMAM* [88] and *ELMAM2* [89] databanks) or based on high-level theoretical computations: The University at Buffalo Pseudoatom Databank (UBDB [90-92]) and the *Generalized Invariom Database* (GID [93, 94]). I will not go into much details about these libraries, their construction and the associated bibliography because Birger Dittrich will discuss this matter in a dedicated lecture in this Erice school. Let's just say all these transferable electron densities present a considerable interest when applied to the study of biological macromolecules. They allow indeed the construction of a continuous and accurate electron density distribution of protein structures, in a matter of minutes. Such transferred charge distribution gives afterward access to the computation of properties that can be directly correlated to the function of the protein, or to the binding mode of a ligand (electrostatic potential, energies, electron density critical points …). Moreover, following the electron density transfer, a structural refinement of a protein structure against subatomic resolution X-ray data allows (i) modeling the residual valence electron density peaks observed after a classical spherical refinement, (ii) improving the structural model, and (iii) reducing the mean square atomic displacements amplitudes. All these applications will be exemplified with the published studies I will discuss in my lecture. With the development of libraries containing transferable electron density fragments emerged the need for a software compatible with the charge density analysis of macromolecular systems. The *MoPro* program package (*MoPro*, *VMoPro* and *MoProViewer*)

filled that void by providing to the charge density community a software able to handle the least-squares refinement of protein structures of reasonable sizes [95,96]. *MoPro* indeed implements the spherical and multipolar atom models as well as features such as for instance bulk solvent modelling and stereochemical and electron density restraints.

Charge density studies of biological macromolecules

All published experimental (in a broad sense of the term) charge density studies of biological macromolecules are based on a transferred charge distribution using either of the three available libraries. However, one can define two main categories, summarized below.

1 – cases where electrostatic properties and interatomic interactions (notably between protein and ligand atoms) were studied using a transferred charge distribution, without any additional electron density refinement.

This approach is undoubtedly the easiest to implement. Indeed, strictly speaking, only the experimental geometry of the protein under exam is needed, even though the structure has not been solved at atomic or subatomic resolution. However, in such cases, the problem of "missing hydrogen atoms" must be addressed: their presence in the atomic model is indeed compulsory to obtain meaningful electrostatic or topological properties, and because the transfer procedure needs all of them in the definition of electron density databases "atom types". Fortunately, missing hydrogen atoms can be added using very efficient tools such as the ones available on the MolProbity web server [97, and references therein]. If needed, their positions can be manually checked / adjusted afterward on the basis of chemical intuition (hydrogen bonds, steric clashes …). Once the model is complete, electron density parameters can be transferred to the studied protein structure, and properties relevant to the protein function can be computed without additional refinement.

This approach has been followed in several studies of proteins or of protein-ligand complexes. The authors discussed (among other things) either properties related to the electrostatic potential [98-101], electron density topological analysis [102, 103], or electrostatic interaction energies between ligand(s) and receptor(s) [104-106]. In the latter case, the molecular electron density decomposition into pseudo-atoms (a property of the multipole modelling), turns out to be especially convenient. It allows indeed to easily decompose the total electrostatic interaction energy into fragments contributions, showing for instance which amino acid have the strongest influence on the ligand binding energy. Electrostatic interaction energies computed from a transferred charge distribution was also exploited to score docking poses of vitamin D analogues in the human vitamin D receptor [107]. The hydrogen atoms issue can be overcome by the use of neutron diffraction on a fully deuterated protein. This has been applied to study the properties of an ordered water cluster in the binding pocket of a Fatty Acid Binding Protein [108]. A room temperature X-ray structure solved at atomic resolution was complemented by a neutron diffraction experiment to model the orientation of water molecules on the basis of their deuterium nuclear densities. A transferred electron density distribution was then used to

characterize the binding of the fatty acid (Fig. 1) and to study the influence of electric field generated by the protein on the orientation of the structured water molecules.

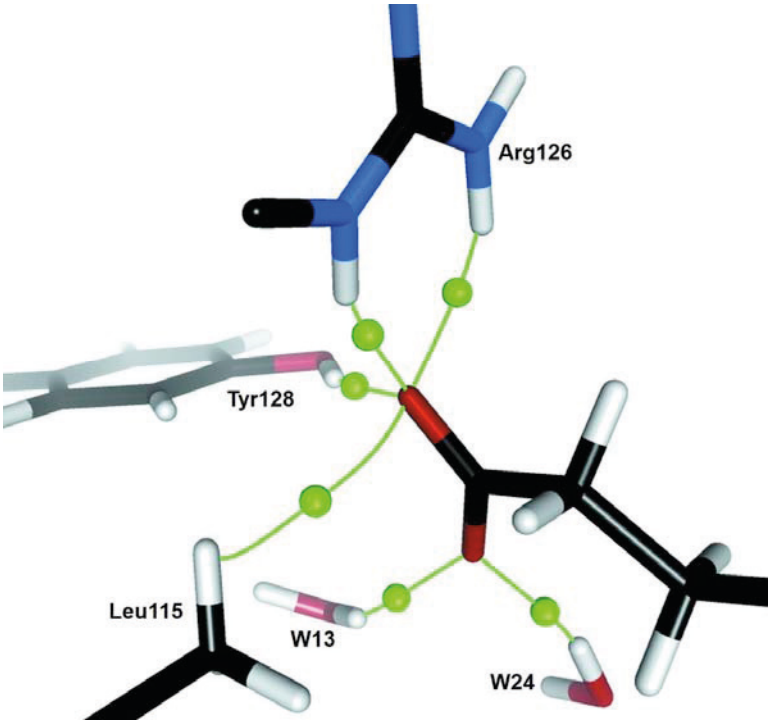


Figure 1. Bond critical points and associated bond paths (pictured in green) of hydrogen bonds involving oxygen atoms of the oleic acid in the FABP binding pocket (figure taken from [108]).

2 – cases where a charge distribution was transferred to a protein model, and the transferred pseudo-atoms were used as a starting point to refine structural parameters, or both structural *and* electron density parameters, before computation of properties.

Several published studies of proteins describe an electron density transfer procedure followed by the refinement of atomic coordinates and thermal displacement parameters [99, 101, 109, 110]. In many cases, authors reported a systematic decrease of atomic B factors, of crystallographic agreement factors (including R_{free}) and, as expected, a flattening of Fourier difference maps. This was also observed after the refinement of Z-DNA hexamer and dodecamer duplexes at respectively 0.55Å and 0.9Å resolution using the UBDB databank [110]. Afterward, electron density derived properties can again be computed and interpreted, such as in the charge density analysis of lysozyme (d = 0.65Å), in which dynamic and static electron density maps were compared, followed by a topological analysis of both densities in the lysozyme active site where the enzymatic mechanism occurs [109].

Finally, there are studies in which refinements of the transferred charge density parameters are also reported [111-117]. Such approaches are trickier as they rely on truly subatomic resolution

X-ray data of sufficiently good quality, as well as on moderate atomic B factors, at least in the most ordered parts of the model. Such requirements are only met in exceptional cases: the decisive criterion being the *actual* observation of significant bonding electron density peaks in residual Fourier maps after the spherical refinement of the protein model. These studies took advantage of the repetition of identical chemical moieties along the protein polypeptide chain, which allowed the use of numerous chemical equivalence constraints in the multipole refinement. In the human aldose reductase charge density analysis [114], methodological issues were discussed, such as the refinement strategy, the use of bulk solvent modelling, of constraints and restraints and of high-order refinement to achieve a partial deconvolution of valence electron density and thermal displacement parameters. A slightly modified refinement strategy was followed by the authors of a recent work [116,117] reporting the charge density analysis of a high-potential iron-sulfur protein (at d = 0.48Å), and the very first experimental electron density characterization in a $[Fe_4S_4]$ cluster relevant to the function of a metalloprotein. To conclude, I suggest again few reviews in this field [118-120].

Bibliography

- **Charge density studies of amino acids, peptides and related derivatives**

[1] Souhassou M, Lecomte C, Blessing RH, Aubry A, Rohmer MM, Wiest R, Bénard M, Marraud M. *Electron distributions in peptides and related molecules. 1. An experimental and theoretical study of N-acetyl-L-tryptophan methylamide*. Acta Crystallogr B. **1991**, 47, 253-66.

[2] Gatti C, Bianchi R, Destro R, Merati F. *Experimental vs. theoretical topological properties of charge density distributions. An application to the L-alanine molecule studied by X-ray diffraction at 23 K* J. Mol. Struct. (THEOCHEM), **1992**, 255, 409-33.

[3] Espinosa E, Lecomte C, Molins E, Veintemillas S, Cousson A, Paulus W. *Electron density study of a new non-linear optical material: L-arginine phosphate monohydrate (LAP). Comparison between X-X and X-(X+N) refinements*. Acta Crystallogr B. **1996**, 52, 519-534.

[4] Flaig R, Koritsanszky T, Zobel D, Luger P. *Topological Analysis of the Experimental Electron Densities of Amino Acids. 1. d,l-Aspartic Acid at 20 K*. J. Am. Chem. Soc.. **1998**, 120(10), 2227-38.

[5] Koritsanszky T, Flaig R, Zobel D, Krane H, Morgenroth W, Luger P. *Accurate experimental electronic properties of dl-proline monohydrate obtained within 1 Day*. Science. **1998**, 279(5349), 356-8.

[6] Flaig R, Koritsánszky T, Janczak J, Krane HG, Morgenroth W, Luger P. *Fast experiments for charge-density determination: Topological analysis and electrostatic potential of the amino acids L-Asn, DL-Glu, DL-Ser, and L-Thr*. Angwewandte Chemie Int. Ed. **1999**, 38(10), 1397-1400.

[7] Coppens P, Abramov Y, Carducci M, Korjov B, Novozhilova I, Alhambra C, Mark R. Pressprich MR. *Experimental Charge Densities and Intermolecular Interactions: Electrostatic and Topological Analysis of dl-Histidine.* J. Am. Chem. Soc. **1999**, 121 (11), 2585–93.

[8] Dahaoui S, Jelsch C, Howard JA, Lecomte C. *Charge density study of N-acetyl-L-tyrosine ethyl ester monohydrate derived from CCD area detector data*. Acta Crystallogr B. **1999**, 55, 226-230.

[9] Dahaoui S, Pichon-Pesme V, Howard JAK, Lecomte C. CCD *Charge Density Study on Crystals with Large Unit Cell Parameters: The Case of Hexagonal l-Cystine*. J. Phys. Chem. A, **1999**, 103 (31), 6240–50.

[10] Dittrich B, Flaig R, Koritsánszky T, Krane HG, Morgenroth W, Luger P. *Topological properties of the peptide bond in glycyl-L-threonine dihydrate based on a fast synchrotron/CCD-diffraction experiment at 100 K*. Chemistry. **2000**, 6(14), 2582-9.
[11] Destro R, Roversi P, Barzaghi M, Marsh RE. *Experimental Charge Density of α Glycine at 23 K.* J. Phys. Chem. A, **2000**, 104, 1047-54.
[12] Benabicha F, Pichon-Pesme V, Jelsch C, Lecomte C, Khmou A. *Experimental charge density and electrostatic potential of glycyl-L-threonine dihydrate*. Acta Crystallogr B. **2000**, 56, 55-65.
[13] Pichon-Pesme V, Lachekar H, Souhassou M, Lecomte C. *Electron density and electrostatic properties of two peptide molecules: tyrosyl-glycyl-glycine monohydrate and glycyl-aspartic acid dihydrate*. Acta Crystallogr B. **2000**, 56, 728-37.
[14] Guillot R, Muzet N, Dahaoui S, Lecomte C, Jelsch C. *Experimental and theoretical charge density of DL-alanyl-methionine*. Acta Crystallogr B. **2001**, 57, 567-78.
[15] Wagner A, Luger P. *Charge density and topological analysis of L-glutamine* J. Mol. Struct. **2001**, 595, 39-46.
[16] Flaig R, Koritsanszky T, Dittrich B, Wagner A, Luger P. *Intra- and Intermolecular Topological Properties of Amino Acids: A Comparative Study of Experimental and Theoretical Results* J. Am. Chem. Soc. **2002**, 124(13), 3407–17.
[17] Li X, Wu G, Abramov YA, Volkov AV, Coppens P. *Application of charge density methods to a protein model compound: Calculation of Coulombic intermolecular interaction energies from the experimental charge density*. PNAS. **2002**, 99(19), 12132-137.
[18] Kingsford-Adaboh R, Dittrich B, Wagner A, Messerschmidt M, Flaig R, Luger P. *Topological analysis of DL-arginine monohydrate at 100 K*. Zeitschrift fur Kristallographie, **2002**, 217(4), 168-173.
[19] Scheins S, Dittrich B, Messerschmidt M, Paulmann C, Luger P. *Atomic volumes and charges in a system with a strong hydrogen bond: L-tryptophan formic acid*. Acta Crystallogr B. **2004**, 60, 184-90.
[20] Dittrich B, Hübschle CB, Messerschmidt M, Kalinowski R, Girnt D, Luger P. *The invariom model and its application: refinement of D,L-serine at different temperatures and resolution*. Acta Crystallogr A. **2005**, 61, 314-20.
[21] Rödel E, Messerschmidt M, Dittrich B, Luger P. *Atomic and bond topological properties of the tripeptide L-alanyl-L-alanyl-L-alanine based on its experimental charge density obtained at 20 K*. Org Biomol Chem. **2006**, 4(3), 475-81.
[22] Mata I, Espinosa E, Molins E, Veintemillas S, Maniukiewicz W, Lecomte C, Cousson A, Paulus W. *Contributions to the application of the transferability principle and the multipolar modeling of H atoms: electron-density study of L-histidinium dihydrogen orthophosphate orthophosphoric acid*. I. Acta Crystallogr. A. **2006**, 62, 365-78.
[23] Checińska L, Mebs S, Hübschle CB, Förster D, Morgenroth W, Luger P. *Reproducibility and transferability of topological data: experimental charge density study of two modifications of L-alanyl-L-tyrosyl-L-alanine*. Org. Biomol Chem. **2006**, 4(17), 3242-51.
[24] Kalinowski R, Dittrich B, Hübschle CB, Paulmann C, Luger P. *Experimental charge density of L-alanyl-L-prolyl-L-alanine hydrate: classical multipole and invariom approach, analysis of intra- and intermolecular topological properties*. Acta Crystallogr. B. **2007**, 63, 753-67.

[25] Hofmann A, Kalinowski R, Luger P, van Smaalen S. *Accurate charge density of the tripeptide Ala-Pro-Ala with the maximum entropy method (MEM): influence of data resolution*. Acta Crystallogr. B. **2007**, 63, 633-43.
[26] Destro R, Soave R, Barzaghi M. *Physicochemical properties of zwitterionic L- and DL-alanine crystals from their experimental and theoretical charge densities*. J Phys Chem B. **2008**,112(16), 5163-74.
[27] Grabowsky S, Kalinowski R, Weber M, Förster D, Paulmann C, Luger P. *Transferability and reproducibility in electron-density studies--bond-topological and atomic properties of tripeptides of the type L-alanyl-X-L-alanine*. Acta Crystallogr B. **2009**, 65, 488-501.
[28] Mebs S, Messerschmidt M, Luger P. *Experimental charge density of an L-phenylalanine formic acid complex with a short hydrogen bond determined at 25 K*. Zeitschrift fur Kristallographie, **2009**, 221(9), 656-664.

- **Charge density studies of nucleic acids and derivatives**

[29] Eisenstein M. *Static deformation densities for cytosine and adenine.* Acta Crystallogr B. **1988**, 44, 412-26.
[30] Weber HP, Craven BM. *Electrostatic properties of cytosine monohydrate from diffraction data*. Acta Crystallogr B. **1990**, 46, 532-8.
[31] Klooster WT, Swaminathan S, Nanni R, Craven BM. *Electrostatic properties of 1-methyluracil from diffraction data*. Acta Crystallogr B. **1992**, 48, 217-27.
[32] Cunane LM, Taylor MR. *Experimental charge density and electrostatic potential in adenine hydrochloride hemihydrate at 123 K*. Acta Crystallogr B. **1993**, 49, 524-30.
[33] Chen L, Craven BM. *Electrostatic properties of beta-cytidine and cytosine monohydrate from Bragg diffraction*. Acta Crystallogr B. **1995**, 51, 1081-96.
[34] Cunane LM, Taylor MR. *The effects of metal binding on a nucleobase: the experimental charge density and electrostatic potential in 1H(+)-adeniniumtrichlorozinc(II) at 123 K and its relationship to that in adenine hydrochloride hemihydrate*. Acta Crystallogr D. **1997**, 53, 765-76.
[35] Guillot B, Muzet N, Artacho E, Lecomte C, Jelsch C. *Experimental and Theoretical Electron Density Studies in Large Molecules: NAD+, β-Nicotinamide Adenine Dinucleotide*. J. Phys. Chem. B. **2003**, 107(34), 9109–21.
[36] Bendeif el-E, Dahaoui S, Benali-Cherif N, Lecomte C. *Tautomerism and hydrogen bonding in guaninium phosphite and guaninium phosphate salts*. Acta Crystallogr B. **2007**, 63, 448-58.
[37] Munshi P, Guru Row TN. *Intra- and intermolecular interactions in small bioactive molecules: cooperative features from experimental and theoretical charge-density analysis*. Acta Crystallogr B. **2006**, 62, 612-26.
[38] Hübschle CB, Dittrich B, Grabowsky S, Messerschmidt M, Luger P. *Comparative experimental electron density and electron localization function study of thymidine based on 20 K X-ray diffraction data*. Acta Crystallogr B. **2008**, 64, 363-74.
[39] Jarzembska KN, Kubsik M, Kamiński R, Woźniak K, Dominiak PM. *From a Single Molecule to Molecular Crystal Architectures: Structural and Energetic Studies of Selected Uracil Derivatives.* Crystal Growth & Design. **2012**, 12(5), 2508-24.

[40] Jarzembska KN, Kamiński R, Wenger E, Lecomte C, Dominiak PM. *Interplay between Charge Density Distribution, Crystal Structure Energetic Features, and Crystal Morphology of 6-Methyl-2-thiouracil*. J. Phys. Chem. C. **2013**, 117(15), 7764-75.
[41] Jarzembska KN, Řlepokura K, Kamiński R, Gutmann MJ, Dominiak PM, Woźniak K. *Multi-temperature study of potassium uridine-5'-monophosphate: electron density distribution and anharmonic motion modelling*. Acta Crystallogr B. **2017**, 73, 550-564.

- **Charge density studies of other endogenous biomolecules: hormones, carbohydrates, neurotransmitters, cofactors**

[42] Wiest R, Pichon-Pesme V, Benard M, Lecomte C. *Electron distributions in peptides and related molecules. Experimental and theoretical study of Leu-enkephalin trihydrate*. J. Phys. Chem. **1994**, 98(4), 1351–62.
[43] Hibbs DE, Austin-Woods CJ, Platts JA, Overgaard J, Turner P. *Experimental and theoretical charge density study of the neurotransmitter taurine*. Chemistry. **2003**, 9(5), 1075-84.
[44] Lo Presti L, Soave R, Destro R. *On the interplay between CH..O and OH...O interactions in determining crystal packing and molecular conformation: an experimental and theoretical charge density study of the fungal secondary metabolite austdiol ($C_{12}H_{12}O_5$)*. J. Phys Chem B. **2006**, 110(12), 6405-14.
[45] Parrish D, Zhurova EA, Kirschbaum K, Pinkerton AA. *Experimental charge density study of estrogens: 17 beta-estradiol-urea*. J. Phys Chem B. **2006**, 110(51), 26442-7.
[46] Zhurova EA, Matta CF, Wu N, Zhurov VV, Pinkerton AA. *Experimental and theoretical electron density study of estrone*. J Am Chem Soc. **2006**, 128, 8849-61.
[47] Yearley EJ, Zhurova EA, Zhurov VV, Pinkerton AA. *Binding of genistein to the estrogen receptor based on an experimental electron density study*. J Am Chem Soc. **2007**, 129(48), 15013-21.
[48] Jaradat DM, Mebs S, Checińska L, Luger P. *Experimental charge density of sucrose at 20K: bond topological, atomic, and intermolecular quantitative properties*. Carbohydr. Res. **2007**, 342(11), 1480-9.
[49] Mebs S, Henn J, Dittrich B, Paulmann C, Luger P. *Electron densities of three B12 vitamins*. J. Phys Chem A. **2009**, 113(29), 8366-78.
[50] Zhurova EA, Zhurov VV, Chopra D, Stash AI, Pinkerton AA. *17-Alpha-estradiol x 1/2 H2O: super-structural ordering, electronic properties, chemical bonding, and biological activity in comparison with other estrogens*. J Am Chem Soc. **2009**, 131(47), 17260-9.
[51] Stevens ED, Dowd MK, Johnson GP, French AD. *Experimental and theoretical electron density distribution of alpha,alpha-trehalose dihydrate*. Carbohydr Res. **2010**, 345(10), 1469-81.
[52] Zhurova EA, Zhurov VV, Kumaradhas P, Cenedese S, Pinkerton AA. *Charge Density and Electrostatic Potential Study of 16α,17β-Estriol and the Binding of Estrogen Molecules to the Estrogen Receptors ERα and ERβ*. J Phys Chem B. **2016** ,120, 8882-91.

- **Charge density studies of pharmaceutical drugs**

[53] Miwa Y, Mizuno T, Tsuchida K, Taga T, Iwata Y. *Experimental charge density and electrostatic potential in nicotinamide*. Acta Crystallogr B. **1999**, 55, 78-84.
[54] Bouhmaida N, Dutheil M, Ghermani NE, Becker P. *Gradient vector field and properties of the experimental electrostatic potential: application to Ibuprofen drug molecule*. J. Chem. Phys. **2002**, 116, 6196–6204.

[55] Wagner A, Flaig R, Dittrich B, Schmidt H, Koritsánszky T, Luger P. *Charge density and experimental electrostatic potentials of two penicillin derivatives*. Chemistry. **2004**, 10(12), 2977-82.
[56] Overgaard J, Hibbs DE. *The experimental electron density in polymorphs A and B of the anti-ulcer drug famotidine*. Acta Crystallogr A. **2004**, 60, 480-7.
[57] Messerschmidt M, Scheins S, Luger P. *Charge density of (-)-strychnine from 100 to 15 K, a comparison of four data sets*. Acta Crystallogr B. **2005**, 61, 115-21.
[58] Scheins S, Messerschmidt M, Luger P. *Submolecular partitioning of morphine hydrate based on its experimental charge density at 25 K*. Acta Crystallogr B. **2005**, 61, 443-8.
[59] Hibbs DE, Overgaard J, Howard ST, Nguyen TH. *Experimental charge density of a potential DHO synthetase inhibitor: dimethyl-trans-2-oxohexahydro-pyrimidine-4,6-dicarboxylate*. Org Biomol Chem. **2005**, 3, 441-7.
[60] Destro R, Soave R, Barzaghi M, Lo Presti L. *Progress in the understanding of drug-receptor interactions, Part 1: experimental charge-density study of an angiotensin II receptor antagonist (C30H30N6O3S) at T = 17 K*. Chemistry. **2005**, 11(16), 4621-34.
[61] Soave R, Barzaghi M, Destro R. *Progress in the understanding of drug-receptor interactions, part 2: experimental and theoretical electrostatic moments and interaction energies of an angiotensin II receptor antagonist (C30H30N6(O)3S)*. Chemistry. **2007**, 13(24), 6942-56.
[62] Overgaard J, Turel I, Hibbs DE. *Experimental electron density study of a complex between copper(II) and the antibacterial quinolone family member ciprofloxacin*. Dalton Trans. **2007**, 21, 2171-8.
[63] Scheins S, Messerschmidt M, Morgenroth W, Paulmann C, Luger P. *Electron density analyses of opioids: a comparative study*. J Phys Chem A. **2007**, 111(25), 5499-508.
[64] Grabowsky S, Pfeuffer T, Morgenroth W, Paulmann C, Schirmeister T, Luger P. *A comparative study on the experimentally derived electron densities of three protease inhibitor model compounds*. Org Biomol Chem. **2008**, 6(13), 2295-307.
[65] Johnas SK, Dittrich B, Meents A, Messerschmidt M, Weckert EF. *Charge-density study on cyclosporine A*. Acta Crystallogr D. **2009**, 65, 284-93.
[66] Bouhmaida N, Bonhomme F, Guillot B, Jelsch C, Ghermani NE. *Charge density and electrostatic potential analyses in paracetamol*. Acta Crystallogr B. **2009**, 65, 363-74.
[67] Dittrich B, Bond CS, Kalinowski,R, Spackman MA, Jayatilaka D. *Revised electrostatics from invariom refinement of the 18-residue peptaibol antibiotic trichotoxin A50E*. Cryst. Eng. Comm. **2010**, 12, 2419-23.
[68] Holstein JJ, Luger P, Kalinowski R, Mebs S, Paulman C, Dittrich B. *Validation of experimental charge densities: refinement of the macrolide antibiotic roxithromycin*. Acta Crystallogr B. **2010**, 66, 568-77.
[69] Weber M, Grabowsky S, Hazra A, Naskar S, Banerjee S, Mondal NB, Luger P. *Electron density of two bioactive oligocyclic indole and oxindole derivatives obtained from low-order X-ray data and invariom application*. Chem Asian J. **2011**, 6(6), 1390-7.
[70] Holstein JJ, Hübschle CB, Dittrich B. *Electrostatic properties of nine fluoroquinolone antibiotics derived directly from their crystal structure refinements*. Cryst. Eng. Comm. **2012**, 14, 2520-31.
[71] Luger P, Weber M, Hübschle CB, Tacle R. *Electron densities of bexarotene and disila-bexarotene from invariom application: a comparative study*. *Org. Biomol. Chem.*, **2013**,11, 2348-54.
[72] Luger P, Weber M, Dittrich B. *Electron density study of the anti-Alzheimer's disease drug donepezil from conventional x-ray data and invariom database application*. Future Med Chem. **2012**, 4, 1399-407.

[73] Luger P, Weber M, Hübschle C, Tacke R. *Electron densities of bexarotene and disila-bexarotene from invariom application: a comparative study*. Org Biomol Chem. **2013**, 11, 2348-54.
[74] Rajalakshmi G, Hathwar VR, Kumaradhas P. *Topological analysis of electron density and the electrostatic properties of isoniazid: an experimental and theoretical study*. Acta Crystallogr B. **2014**, 70, 331-41.
[75] Du JJ, Hanrahan JR, Solomon VR, Williams PA, Groundwater PW, Overgaard J, Platts JA, Hibbs DE. *Exploring the Binding of Barbital to a Synthetic Macrocyclic Receptor. A Charge Density Study*. J Phys Chem A. **2018,** 122(11), 3031-44.

- **Reviews on charge density in medicine, drug design …**

[76] Luger P. *Fast electron density methods in the life sciences--a routine application in the future?* Org Biomol Chem. **2007**, 5(16), 2529-40.
[77] Matta CF, Arabi AA. *Electron-density descriptors as predictors in quantitative structure--activity/property relationships and drug design*. Future Med Chem. **2011**, 3(8), 969-94.
[78] Dittrich B, Matta CF. *Contributions of charge-density research to medicinal chemistry*. IUCrJ. **2014**, 1, 457-69.
[79] Matta CF. *Modeling biophysical and biological properties from the characteristics of the molecular electron density, electron localization and delocalization matrices, and the electrostatic potential*. J. Comput. Chem. **2014**, 35(16), 1165-98.

- **Methods (bio crystallography, electron density databases, computer programs …)**

[80] Giegé R, Sauter C. *Biocrystallography: past, present, future*. HFSP Journal. **2010**, 41, 109-121.
[81] Chruszcz M, Potrzebowski W, Zimmerman MD, Grabowski M, Zheng H, Lasota P, Minor W. *Analysis of solvent content and oligomeric states in protein crystals—does symmetry matter?* Protein Science: A Publication of the Protein Society, **2008**, *17*(4), 623–632.
[82] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. *The Protein Data Bank*. Nucleic Acids Research, **2000**, 28, 235-242.
[83] Brock CP, Dunitz JD, Hirshfeld FL. *Transferability of deformation densities among related molecules: atomic multipole parameters from perylene for improved estimation of molecular vibrations in naphthalene and anthracene*. Acta Crystallogr B. **1991**, 47, 789-97.
[84] Pichon-Pesme V, Lecomte C, Lachekar H. *On Building a Data Bank of Transferable Experimental Electron Density Parameters Applicable to Polypeptides*. J. Phys. Chem. **1995**, 99, 6242–50.
[85] Jelsch C, Pichon-Pesme V, Lecomte C, Aubry A. *Transferability of multipole charge-density parameters: application to very high resolution oligopeptide and protein structures*. Acta Crystallogr D. **1998**, 54, 1306-18.
[86] Pichon-Pesme V, Jelsch C, Guillot B, Lecomte C. *A comparison between experimental and theoretical aspherical-atom scattering factors for charge-density refinement of large molecules.* Acta Crystallogr. A. ***2004**, 60, 204-8.*
[87] Hansen NK, Coppens P. *Testing aspherical atom refinements on small-molecule data sets*. Acta Crystallogr A, **1978**, 34, 909–921
[88] Zarychta B, Pichon-Pesme V, Guillot B, Lecomte C, Jelsch C. *On the application of an experimental multipolar pseudo-atom library for accurate refinement of small-molecule and protein crystal structures*. Acta Crystallogr A. **2007**, 63, 108-25.

[89] Domagała S, Fournier B, Liebschner D, Guillot B, Jelsch C. *An improved experimental databank of transferable multipolar atom models--ELMAM2. Construction details and applications.* Acta Crystallogr A. **2012,** 68, 337-51.
[90] Volkov A, Li X, Koritsanszky T, Coppens P. *Ab Initio Quality Electrostatic Atomic and Molecular Properties Including Intermolecular Energies from a Transferable Theoretical Pseudoatom Databank*. J. Phys. Chem. A. **2004**, 108(19), 4283-4300.
[91] Dominiak PM, Volkov A, Li X, Messerschmidt M, Coppens P. A *Theoretical Databank of Transferable Aspherical Atoms and Its Application to Electrostatic Interaction Energy Calculations of Macromolecules.* J Chem Theory Comput. **2007,** 3(1), 232-47.
[92] Jarzembska KN, Dominiak PM. *New version of the theoretical databank of transferable aspherical pseudoatoms, UBDB2011--towards nucleic acid modelling.* Acta Crystallogr A. **2012**, 68, 139-47.
[93] Dittrich B, Hübschle CB, Luger P, Spackman MA. *Introduction and validation of an invariom database for amino-acid, peptide and protein molecules.* Acta Crystallogr. D, **2006**, 1325-35.
[94] Dittrich B, Hübschle CB, Pröpper K, Dietrich F, Stolper T, Holstein JJ. *The generalized invariom database (GID).* Acta Crystallogr B. **2013**, 69, 91-104.
[95] Guillot B, Viry L, Guillot R, Lecomte C, Jelsch C. *Refinement of proteins at subatomic resolution with MoPro*. J. App. Cryst. **2001**, 34, 214-23.
[96] Jelsch C, Guillot B, Lagoutte A, Lecomte C. *Advances in protein and small-molecule charge-density refinement methods using MoPro*. J. App. Cryst. **2005**, 38, 38-54.
[97] Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB 3rd, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. *MolProbity: More and better reference data for improved all-atom structure validation*. Protein Sci. **2018**, 27, 293-315.

- **Charge density studies of proteins.**

[98] Muzet N, Guillot B, Jelsch C, Howard E, Lecomte C. *Electrostatic complementarity in an aldose reductase complex from ultra-high-resolution crystallography and first principles calculations*. PNAS, **2003**, 100(15), 8742-7.
[99] Zarychta B, Lyubimov A, Ahmed M, Munshi P, Guillot B, Vrielink A, Jelsch C. *Cholesterol oxidase: ultrahigh-resolution crystal structure and multipolar atom model-based analysis*. Acta Crystallogr D. **2015**, 71, 954-68.
[100] Liebschner D, Elias M, Moniot S, Fournier B, Scott K, Jelsch C, Guillot B, Lecomte C, Chabrière E. *Elucidation of the Phosphate Binding Mode of DING Proteins Revealed by Subangstrom X-ray Crystallography*. J. Am. Chem. Soc. **2009**, 131, 7879-86.
[101] Pröpper K, Holstein JJ, Hübschle CB, Bond CS, Dittrich B. *Invariom refinement of a new monoclinic solvate of thiostrepton at 0.64 Å resolution*. Acta Crystallogr D. **2013**, 69, 1530-9.
[102] Liebschner D, Jelsch C, Espinosa E, Lecomte C, Chabrière E, Guillot B. *Topological analysis of hydrogen bonds and weak interactions in protein helices via transferred experimental charge density parameters.* J Phys Chem A. **2011**, 115(45), 12895-904.
[103] Takaba K, Takeda K, Kosugi M, Tamada T, Miki K. *Distribution of valence electrons of the flavin cofactor in NADH-cytochrome b(5) reductase*. Sci Rep. **2017**, 7, 43162.

[104] Dominiak PM, Volkov A, Dominiak AP, Jarzembska KN, Coppens P. *Combining crystallographic information and an aspherical-atom data bank in the evaluation of the electrostatic interaction energy in an enzyme-substrate complex: influenza neuraminidase inhibition*. Acta Crystallogr. D. **2009**, 65, 485-99.
[105] Malińska M, Jarzembska KN, Goral AM, Kutner A, Woźniak K, Dominiak PM. *Sunitinib: from charge-density studies to interaction with proteins*. Acta Crystallogr D. **2014**, 70, 1257-70.
[106] Fournier B, Bendeif el-E, Guillot B, Podjarny A, Lecomte C, Jelsch C. *Charge density and electrostatic interactions of fidarestat, an inhibitor of human aldose reductase*. J Am Chem Soc. **2009**, 131(31), 10929-41.
[107] Malinska M, Kutner A, Woźniak K. *Predicted structures of new Vitamin D Receptor agonists based on available X-ray structures*. Steroids. **2015**, 104, 220-9.
[108] Howard EI, Guillot B, Blakeley MP, Haertlein M, Moulin M, Mitschler A, Cousido-Siah A, Fadel F, Valsecchi WM, Tomizaki T, Petrova T, Claudot J, Podjarny A. *High-resolution neutron and X-ray diffraction room-temperature studies of an H-FABP-oleic acid complex: study of the internal water cluster and ligand binding by a transferred multipolar electron-density distribution*. IUCrJ. **2016**, 3, 115-26.
[109] Held J, van Smaalen S. *The active site of hen egg-white lysozyme: flexibility and chemical bonding.* Acta Crystallogr D, **2014**, 70, 1136-1146.
[110] Malinska M, Dauter Z. *Transferable aspherical atom model refinement of protein and DNA structures against ultrahigh-resolution X-ray data*. Acta Crystallogr D. **2016**, 72, 770-9.
[111] Housset D, Benabicha F, Pichon-Pesme V, Jelsch C, Maierhofer A, David S, Fontecilla-Camps JC, Lecomte C. *Towards the charge-density study of proteins: a room-temperature scorpion-toxin structure at 0.96 A resolution as a first test case*. Acta Crystallogr D. **2000**, 56, 151-60.
[112] Jelsch C, Teeter MM, Lamzin V, Pichon-Pesme V, Blessing RH, Lecomte C. *Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin*. PNAS. **2000**, 97(7), 3171-6.
[113] Schmidt A, Jelsch C, Ostergaard P, Rypniewski W, Lamzin VS. *Trypsin revisited: crystallography AT (SUB) atomic resolution and quantum chemistry revealing details of catalysis*. J. Biol. Chem. **2003**, 278(44), 43357-62.
[114] Guillot B, Jelsch C, Podjarny A, Lecomte C. *Charge-density analysis of a protein structure at subatomic resolution: the human aldose reductase case*. Acta Crystallogr D, **2008**, 64, 567-88.
[115] Schmidt A, Teeter M, Weckert E, Lamzin VS. *Crystal structure of small protein crambin at 0.48 Å resolution*. Acta Crystallogr. F, **2011**, 67, 424-28.
[116] Hirano Y, Takeda K, Miki K. *Charge-density analysis of an iron-sulfur protein at an ultra-high resolution of 0.48 Å*. Nature. **2016**, 534(7606), 281-4.
[117] Takeda K, Miki K. *Ultra-high-resolution structure and charge-density analysis of high-potential iron-sulfur protein*. FEBS J. **2017**, 284(14), 2163-66.
[118] Lecomte C, Guillot B, Muzet N, Pichon-Pesme V, Jelsch C. *Ultra-high-resolution X-ray structure of proteins*. Cell Mol Life Sci. **2004**, 61, 774-82.
[119] Lecomte C, Guillot B, Jelsch C, Podjarny A. *Frontier example in experimental charge density research: Experimental electrostatics of proteins*. Int. J. Quantum Chem. **2005**, 101, 624-34.
[120] Lecomte C, Jelsch C, Guillot B, Fournier B, Lagoutte A. *Ultrahigh-resolution crystallography and related electron density and electrostatic properties in proteins*. J Synchrotron Radiat. **2008**, 15, 202-3.

Organometallic bonding (concepts) under pressure

Wolfgang Scherer
Institute of Physics, University of Augsburg, Augsburg, Germany
wolfgang.scherer@physik.uni-augsburg.de

1. Introduction: The organometallic bond – definition and functionality.

Geoffrey Wilkinson addressed the importance of the organometallic bond in 1974 by the simple fact: "*During the time taken to deliver this lecture, many thousands, if not tens of thousands, of tons of chemical compounds are being transformed or synthesized industrially in processes which at some stage involve a transition metal to carbon bond. The nonchemist will probably be most familiar with polyethylene or polypropylene in the form of domestic utensils, packaging materials, children's toys, and so on. These materials are made by Ziegler-Natta or Phillips' catalysis, which utilize compounds of the metals titanium and chromium, respectively*" (Wilkinson 1974). Besides its unique importance in industrial applications also evolution acknowledged the unique properties of the organometallic bond which occurs in biological systems whenever chemical processes can only be accomplished in an efficient way by organometallic catalysis (e.g in Vitamine B_{12}).

The aim of this lecture is to outline how we can control the kinetic and thermodynamic characteristics of M-C bonds to activate/break M-C bonds (e.g in Eq.1) or to perform insertion reactions (e.g in Eq. 2) in catalytical conversions of industrial relevance.

$M\text{-}CR_3 + H_2 \rightarrow M\text{-}H + HCR_3$ (1)

$M\text{-}CR_3 + CO \rightarrow M\text{-}CO\text{-}CR_3$ (2)

As a preparation of this lecture I will provide you in the following some basic information concerning of the bonding characteristics of M-C bonds in metal organyls.

In general, we consider compounds with direct metal to carbon bonds as *organometallic* molecules or solids. In this respect they differ from *metal organic* compounds where organic ligands are connected to the metal center by other non-metallic elements such as oxygen (e.g. metal alkoxides) or nitrogen (e.g. metal amides; Elchenbroich 2006). Compounds involving only carbon and metal to establish the organometallic M-C bond are classified as *metal carbides* which are often characterized by puzzling electronic structures and physical properties (Scherer *et al.*, 2012).

The *M*-C bond is usually described as a hetero polar $M^{\delta+}$-$C^{\delta-}$ covalent σ-bond which however can gain multiple bond character e.g. in transition metal *carbene* (M=C) or *carbyne* (M≡C) complexes. When *M* represents a *main group metal* the bonding properties are mainly characterized by the *electronegativity* (EN) differences between the *s,p* block metal and the

neighboring carbon atom of the organic ligand. In the latter case the carbon's electronegativity mainly depends on its hybridization. According *Bent's rules* (Bent, 1961) the EN(C) values increase with increasing *s*-character in the respective hybridorbital: EN(Csp^3) = 2.5, EN(Csp^2) = 2.75 (comparable with the EN of S) und EN(Csp)= 3.29 (comparable with the EN of Cl). Fine tuning of the M-C properties by ligand tailoring can usually be accounted for by determination of group electronegativites (e.g. EN(CH_3) = 2.31 and EN(CF_3) = 3.47 (Bratsch, 1985). However, a large variety of definition of electronegativities exist and identification of chemical trends to classify M-C bonds often depends on the clever choice of suitable EN concepts. This is an unsatisfying situation and we will learn in this lecture that analysis of the *topology of the charge density* provides a reliable and consistent alternative to classify the nature of M-C bonds in organometallic complexes and carbides. Further information can be retrieved experimentally by analyzing the response of M-C bonds under external pressure by *high-pressure diffraction studies*.

2. The nature of the M-C bond in organometallic transition metal complexes

2.1 Kinetic and thermodynamic stability

As outlined in the Intro the M-C bond is usually characterized as a polar $M^{\delta+}$-$C^{\delta-}$ σ-bond. In comparison with the strength of M-N, M-O and M-Cl the M-C bond is rather weak in line Pauling's (1960) classical equation (3) for polar *M*-C dissociation energies *D*. This equation considers the additional the additional electrostatic contribution arising from the electronegativity difference ($\chi_M - \chi_Y$) of the bonding partners in covalent M-Y bonds which increases from Group IV to VII elements *Y*.

$$D_{M-Y} = (D_{M-M} \cdot D_{M-Y})^{0.5} + k|\chi_M - \chi_Y| \qquad (3)$$

However, this simple concept cannot be employed in case of transition metal organyls since the strength of the M-C bonds are usually not controlled by their *thermodynamic dissociation energies* but rather their *kinetic decomposition mechanism* such as the -elimination process rendering a transition metal alkyl in an olefin-hydride species according to Eq. 4 (Wilkinson, 1974).

$$M\text{-}CH_2\text{-}CH_3 \rightarrow M\text{-}H + CH_2{=}CH_2 \qquad (4)$$

Also the stability of oxidation states might control the stability of M-C bonds. For example $Ti(CH_3)_4$ **1** displays a rather larger M-C force constants of 228 Nm^{-1} (Eysel 1970) but readily decomposes above -70°C by an intermolecular decomposition mechanism to from a stable Ti^{III} decomposition product with titanium in a lower oxidation state.

The increasing stability of higher oxidations states down the group of transition metal compounds is therefore reflected in the dissociation energies as illustrated in case of the Group IV metallocene complexes $(Cp^*)_2MMe_2$ (**2-*M***; *M* = Ti,Zr,Hf). Here, the largest bond enthalpy in the case of the Hf–C (306 (7) kJ/mol) compared to the Zr–C (284 (2) kJ/mol) and the Ti–C bond (281 (8) kJ/mol) (Simoes, 1990).

We will now outline that above sequence of bond strength (Hf-C > Zr-C > Ti-C) is also reflected by charge density properties. Due to the inherent instability of the $(Cp^*)_2MMe_2$ complexes we substituted the Cp* ligands ($Cp^* = C_5(Me)_5$) by the sterically more congesting ansa-bridging ligand *L* (L= *rac-Dimethylsilanediylbis[4-(3',5'-dimethylphenyl)-7-methoxy-2-methylindenyl)* (Machat, 2018). Figure 1 shows the corresponding L_2TiMe_2 complex **3-Ti** which is isotypic to the corresponding complexes **3-Zr** and **3-Hf**.

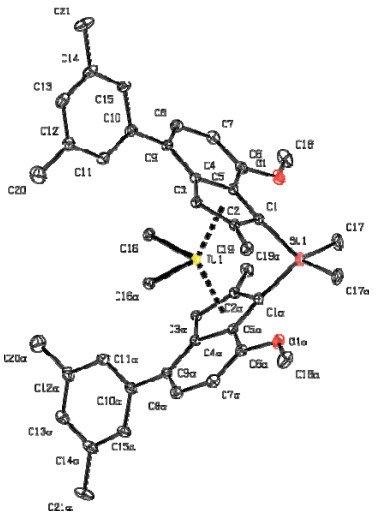


Figure 1. ORTEP style representation of 3-Ti with ellipsoids drawn at 50 % probability level. Hydrogen atoms are omitted for clarity.

2.2 Bond distances

Comparison of the M-C bond distances in these Group IV metallocenes we see that these transition metal organyls clearly reflect systematic trends: **3-Ti**: 2.1527(3) Å; **3-Zr**: 2.2654(3) Å; **3-Hf**: 2.2443(5). The relatively large Ti-C bond in 1-Ti signals a relative weak M-C bond since the 3d metal orbitals which display *no radial node* are poorly *shielded* and thus too *contracted* to form strong covalent M-C bonds. The small differences in the M-C bonds of **3-Zr** and **3-Hf**, however, are a matter of the lanthanoid contraction which dominates the M-C distances of the 5d and 6d metals. The *lanthanoid contraction* is mainly caused by the poor shielding of the 4f electrons in the 6d metals yielding rather small atomic and ionic radii which approximately have the same size as the corresponding 5d metals. To a far lesser extend the lanthanoid contraction is caused by *relativistic effects* (Huheey, 2006).

2.3 Electron Density Characteristics of M-C bonds

We first recall some peculiarities of the *M*–C bond in early d^0 transition metal alkyls. In general, these are considered as highly polar. Indeed, a *partitioning of the experimental electron density* in the framework of the *Quantum Theory of Atoms In Molecules (QTAIM)* (Bader, 1994) supports

the *carbanionic* nature of the methyl carbon atom Q(C) = −0.77 [−0.32] (**3-Ti**), −1.19 [−0.39] (**3-Zr**) and −1.22 [−0.41] (**3-Hf**). Note, that we specify here and in the following values obtained from *density functional theory (DFT)* in square brackets. Due to the pronounced $M \rightarrow$C charge transfer also the methyl hydrogen atoms of **3-*M*** are characterized by *hydridic* (*high-field shifts*) of the corresponding protons in the ^1H *NMR spectra* (δ_{1H} = −0.83 ppm (**3-Ti**), −1.17 ppm (**3-Zr**) and −1.35 ppm (**3-Hf**). Hence, these results suggest an increase of the ionic character of the M–C bond in the sequence **3-Ti** < **3-Zr** < **3-Hf**. This is also evident from the fine structure of the *negative Laplacian* of the electron density, $L(\mathbf{r}) = -\nabla^2\rho(\mathbf{r})$, in the molecular (C,*M*,C) plane (Figure 2). The red solid/blue broken contour lines reveal regions in Figure 2b and c where the electron density is *locally concentrated (L(r) >0) or depleted (L(r) <0)*, respectively (Bader, 1984). Apparently, each methyl carbon atom displays a *bonded charge concentration* (denoted *BCC* in Figure 2b,c) in its valence shell which is located on the M–C bond path and provides a measure of the *lone pair character* at the respective carbon atom (Scherer, 2001, 2002) The magnitude of these BCCs increases in the sequence **3-Ti** < **3-Zr** < **3-Hf** from $L(\mathbf{r})$ = 26.8 − 33.2 eÅ^{-5} in line with the increasing carbanionic character of the respective methyl carbon atoms. However, the $L(\mathbf{r})$ topology reveals also a significant residual covalent character of the M–C bonds despite their polar nature. Especially, in **3-Ti** the covalent nature of the M–C bonds is clearly signaled by the characteristic pattern of two opposing BCCs in the valence shell of metal and the methyl carbon atom (Figure 2b,c) which are connected by a *bond path*.

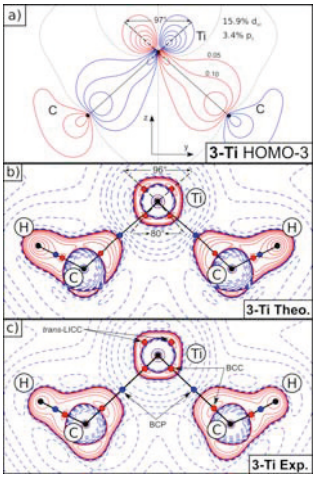


Figure 2. (a) Density contour map of the most important M–C(Me) bonding molecular orbital (HOMO-3) in the molecular plane of **3-Ti**. Contour levels are specified in atomic units and the percentage $\%p_y$ and $\%d_{yz}$ character of the titanium atom refers to gross populations of symmetrized fragment (valence) orbitals (SFOs). Note that the angle between the density-maxima of ligand-opposed p_y-d_{yz} hybrid lobes and the enclosed metal is larger than 90°; (b,c) The topology of the negative Laplacian, $L(r) = -\nabla^2\rho(r)$, reveals ligand-induced charge concentrations in the valence shell of the titanium atom opposite to the methyl ligands (denoted *trans*-LICCs) and pronounced bonded charge concentrations (BCC) in the valence shell of the carbon atom reflecting their carbanionic character. Positive (red, solid) and negative (blue, dashed) $L(r)$ contour lines were drawn at $\pm 2.0^{\times}10^{n}$, $\pm 4.0^{\times}10^{n}$, $\pm 8.0^{\times}10^{n}$ e/Å^5 with n = ±2, ±1, 0.

The formation of pronounced *ligand-induced charge concentrations* (trans-LICCs) opposite to the *M*–C bonds provides another signature of covalent *M*–C bonding (Scherer 2003) and is a natural consequence of the overlap between the sp^x hybridized orbitals at the methyl carbon atoms and Ti($p_y d_{yz}$) hybrids (Figure 2a). The *M*–C covalency is also witnessed by the surprisingly small C–*M*–C angle of 92.79° in **3-Ti**. On contrast the Cp–Ti–Cp angles which are formed by the centroid of the Cp-rings and the metal are significantly larger (132.87°).

To summarize, the increasing ionic character of the *M*–C bond from **3-Ti** to **3-Hf** is clearly supported by the trends observed for the (*i*) atomic charges at the metal and methyl group atoms, (*ii*) the carbanionic character of the carbon atoms, and (*iii*) the hydridic ^{1}H chemical shifts. However, the presence of (*iv*) bonded and ligand opposed charge concentrations (denoted BCC and *trans*-LICCs in Figure 2) in the valence shell of the metal atoms in **3-Ti** and **3-Zr** suggests that the *M*–C bonding scenario is best described by highly polar bonds with residual but decreasing covalent character in the sequence **3-Ti** > **3-Zr** > **3-Hf**.

2.4 Probing the M-C Characteristics by High-Pressure Studies

In the following we have studied the *pressure-induced structural changes* in single crystals of **3-*M***. In general it is assumed that "*forces in crystals act through localized bonds*" which in turn can be identified by the QTAIM methods via the presence of a bond path (Brown, 2003). In case of highly symmetrical covalent compounds, *hardness* is defined by the *bulk modulus B* and controlled by the compressibility of their individual chemical bonds (Liu 1989). These case studies showed that the *bond compressibility* in turn increases with enhancement of the bond ionicity and lowering of the bond strength. In case of molecular crystals, however, the scenario is more complex and the low bulk modulus is often reflecting the softness of *van der Waals contacts* at lower pressure (Brown 2003). It is therefore essential to compare the compressibility of chemical bonds in molecular crystals displaying virtually identical *crystal packing forces*. This is actually the case for **3-Zr** and **3-Hf** which are isostructural and characterized by virtually identical lattice parameters. As a consequence of the lanthanide contraction also the corresponding metal to ligand bonds are very similar (Table 1). Indeed, the Hf-C bond length of 2.2443(5) Å differs by less than 1% from the Zr-C bond distance of 2.2654(3) Å.

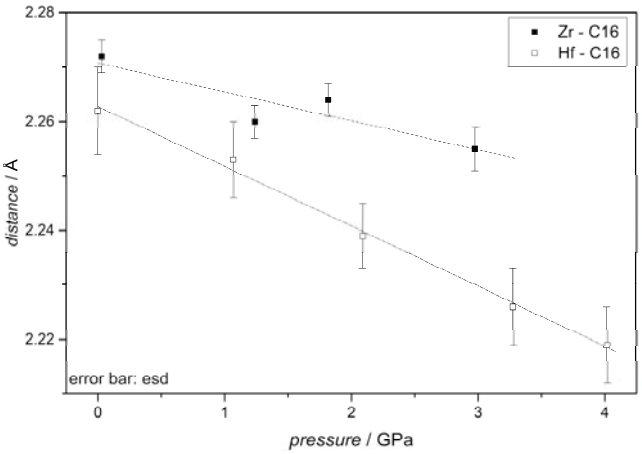


Figure 3. X-ray single-crystal diffraction data of **3-Zr** and **3-Hf** at variable pressure (0 – 4.0(1) GPa). The Hf–C bond shows a larger compressibility.

As a consequence, we expect in case of the σ-*M*–C bonded methyl groups in **3-Zr** and **3-Hf** an increasing bond compressibility with increasing ionic character of the respective *M*–C bonds – if we assume that the bond strength of Hf–C is not significantly larger than that of Zr–C. In that case, the Hf–C bond should display the larger bond compressibility in comparison to the Zr–C bond due to its higher ionic character as determined in the experimental and theoretical charge density studies (see above). Indeed, the experimental high-pressure data (Figure 3) clearly reveal the larger compressibility of the Hf–C bond vs the Zr–C bond. However, analysis of the *potential energy surface* (*PES*) of the individual *M*-C bonds reveals that the Hf-C bond is stiffer than the Zr-C bond. The estimated *force constants* of the *isolated v(M–C) stretching frequencies*, k_{is}, increase down the row of the Group 4 complexes k_{is}/Nm^{-1} = 150 (**3-Ti**), 149 (**3-Zr**) and 168 (**3-Hf**) with a subtle minimum in case of **3-Zr.** This is in line with the cross correlation between the increasing polarity of a chemical bond and the enhancement of its force constant (Haaland 2007). However, we also need to explain why the stiffer Hf–C bond is more compressible compared to the softer Zr-C bond in the **3-*M*** complexes. One might argue empirically that the *bond ionicity rather than the bond strength* predominantly controls the compressibility of the *M*–C bond. However, we will outline in the presentation on Saturday, June 9, more specifically the *microscopic control parameters* of the *M*–C bond compressibilities in transition metal organyls complexes.

References

Bader, R. F. W., MacDougall, P. J., & Lau, C. D. H. (1984). *J. Am. Chem. Soc.* **106**, 1594-1605.

Bader, R. F. W., (1994). *Atoms in Molecules - A Quantum Theory*, Clarendon Press Oxford.

Bent, H. A. (1961). *Chem. Rev*., **61**, 275-311.

Bratsch, S. G. (1985). *J. Chem. Educ.* **62**, 101-103.

Brown, I. D., Klages, P., & Skowron, A. (2003) *Acta Cryst. B*, **59**, 439-448.

Elchenbroich, Ch. (2006) *Organometallics*, Wiley-VCH.

Eysel, H. H., Siebert, H., Groh, G., & Berthold, H. J. (1970). *Spectrochim. Acta* **26A**, 1595-1601.

Haaland, A., (2007). *Molecules and Models*, Oxford, p. 69.

Huheey, J. E., Keiter, E. A., Keiter, R. L. & Medhi O. K. (2006). *Inorganic chemistry: principles of structure and reactivity*, Pearson Education India.

Liu, A. Y., & Cohen, M. L. (1989), *Science*, **245**, 841-842.

Machat, M. R., Fischer, A., Schmitz, D., Vöst, M., Drees, M., Jandl, Ch., Pöthig, A., Casati, N. P. M., Scherer, W., & Rieger, B. (2018), unpublished results.

Pauling, L., 1960. *Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, Chapter 3.

Scherer, W., Sirsch, P., Grosche, M., Spiegler, M., Mason, S. A., & Gardiner, M. G., (2001). *Chemical Commun.*, 2072-2073.

Scherer, W., Sirsch, P., Shorokhov, D., McGrady, G. S., Mason, S. A., & Gardiner, M. G. (2002), *Chem. Eur. J.*, **8**, 2324-2334.

Scherer, W., Sirsch, P., Shorokhov, D., Tafipolsky, M., McGrady, G. S., & Gullo, E., (2003). *Chem. Eur. J.*, **9**, 6057-6070.

Scherer, W., Eickerling, G., Hauf, Ch., Presnitz, M., Scheidt, E.-W., Eyert, V. & Pöttgen, R. (2012) *On the Interplay Between Real and Reciprocal Space Properties, In Modern Charge Density Analysis*, Editors Gatti C., Macchi P. Springer, Netherlands, pp. 359-385.

Wilkinson, G. (1974). *Science*, **185**, 109-112.

Simoes, J. A. M., & Beauchamp, J. L., (1990). *Chem. Rev.* **90**, 629-688.

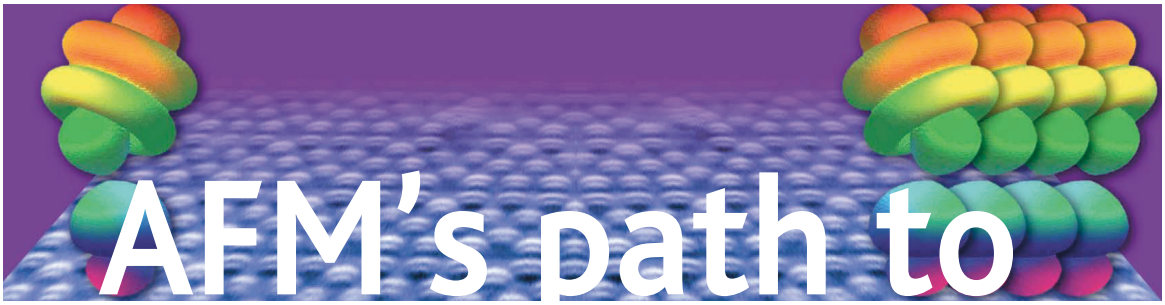
Atomic force microscopy – microscopy with ultimate resolution

Franz J. Giessibl
University of Regensburg, Department of Experimental and Applied Physics, Regensburg, Germany
franz.giessibl@ur.de

Atomic force microscopy (AFM) is a tool that provides, to my knowledge, the best spatial resolution of all microscopy techniques. The lecture will provide a view on this topic from my personal perspective. AFM was introduced by Binnig, Quate and Gerber in 1986 [1] and has evolved into a technique with very broad applications. It has been a great privilege to serve on the improvement of this technique for 30 years now and I hope to bring some of the adventures into the lecture hall. My journey into AFM started as a graduate student with Gerd Binnig in 1988, working on a low-temperature atomic force microscope for vacuum that provided atomic resolution on a chemically inert sample [2] to achieving atomic resolution on the silicon (111) surface as a R&D scientist in a startup company in silicon valley [3]. Frequency modulation AFM is the most precise method to probe surfaces at low noise, therefore taking a loan from the watch industry by transforming the quartz tuning fork of an electronic watch into a force sensing cantilever opened a new level of precision and resolution. In particular, the finding of Gross and Meyer to terminate the tip of the AFM sensor with a CO molecule has brought the possibility to probe organic molecules at atomic resolution [4]. Today, breakthroughs that became available with these new imaging possibilities are reported almost monthly in Science, Nature, Physical Review Letters and other influential journals.

As a preparation for the lecture, I suggest to study a short review [5], or if time permits, a longer one [6].

[1] G. Binnig, C. F. Quate, C. Gerber, Phys. Rev. Lett. (1986)
[2] F. J. Giessibl, G. Binnig, Ultramicroscopy 42, 281 (1992)
[3] F. J. Giessibl, Science 267, 68 (1995).
[4] L. Gross et al., Science 325, 1110 (2009).
[5] F. J. Giessibl, Materials Today 8, 32 (2005).
[6] F. J. Giessibl, Rev. Mod. Phys. 75, 949 (2003).



AFM's path to atomic resolution

by Franz J. Giessibl

We review progress in improving the spatial resolution of atomic force microscopy (AFM) under vacuum. After an introduction to the basic imaging principle and a conceptual comparison to scanning tunneling microscopy (STM), we outline the main challenges of AFM as well as the solutions that have evolved in the first 20 years of its existence. Some crucial steps along AFM's path toward higher resolution are discussed, followed by an outlook on current and future applications.

Experimentalphysik VI, EKM,
Institute of Physics, Augsburg University,
86135 Augsburg, Germany
E-mail: franz.giessibl@physik.uni-augsburg.de

AFM, invented[1] and introduced[2] in 1986, can be viewed as a mechanical profiling technique that generates three-dimensional maps of surfaces by scanning a sharp probe attached to a cantilever over a surface. The forces that act between the tip of the cantilever and the sample are used to control the vertical distance. AFM's potential to reach atomic resolution was foreseen in the original scientific publication[2] but, for a long time, the spatial resolution of AFM was inferior to the resolution capability of its parent technique, STM.

The resolution limits of STM and AFM are given by the structural properties of the atomic wavefunctions of the probe tip and the sample. STM is sensitive to the most loosely bonded electrons with an energy at the Fermi level, while AFM responds to all electrons, including core electrons. Because electrons at the Fermi level are spatially less confined than core electrons, in theory AFM should be able to achieve even greater spatial resolution than STM. Today, experimental evidence is emerging where, in simultaneous AFM/STM studies, AFM images reveal even finer structural details than simultaneously recorded STM images.

The experimental advances that made high-resolution AFM possible began with the introduction of frequency-modulation AFM (FM-AFM). Here, the cantilever oscillates at a fixed amplitude and frequency is used as a feedback signal. Early implementations of FM-AFM used Si cantilevers with a typical spring constant of 10 N/m, which oscillate with an amplitude on the order of 10 nm. The spatial resolution was increased by the introduction of quartz cantilevers with a

ISSN:1369 7021 © Elsevier Ltd 2005

stiffness on the order of 1 kN/m, allowing the use of subnanometer amplitudes. The direct evaluation of higher harmonics in the cantilever motion has enabled a further increase in spatial resolution.

Because AFM can image insulators as well as conductors, it is now a powerful complement to STM for atomically resolved surface studies. Immediate applications of high-resolution AFM have been demonstrated in vacuum studies relating to materials science, surface physics, and surface chemistry. Some of the techniques developed for ultrahigh-vacuum AFM may be applicable for increasing AFM resolution in the ambient or liquid environments that are necessary for studying biological or technological specimens.

Principles of AFM operation

AFM[1,2] can be viewed as an extension of the toddler's way of 'grasping' the world by touching and feeling, as indicated in Fig. 1 of Binnig and Rohrer's article[3], where a finger profiles an atomic surface. Likewise, one could argue that stylus profilometry is a predecessor of AFM. However, AFM and stylus profilometry have as much in common as a candle and a laser. Both generate light and, even though candles are masterpieces of engineering[4], the laser is a much more advanced technological device requiring a detailed knowledge of modern quantum mechanics[5]. While stylus profilometry is an extension of human capabilities that have been known for ages and works by classical mechanics, AFM requires a detailed understanding of the physics of chemical bonding forces and the technological prowess to measure forces that are several orders of magnitude smaller than the forces acting in profilometry.

Only the spectacular spatial resolution of STM could trigger the hope that the force acting between an STM tip and a sample might lead to AFM capable of true atomic resolution. Established in 1981, the STM was the first instrument to allow surface imaging with atomic resolution in real space[6,7]. The atomic imaging of the 7×7 reconstruction of Si (111) by STM in 1983[8] later helped to solve one of the most intriguing problems of surface science at that time and establish the dimer-adatom-stacking fault model of Takayanagi *et al.*[9]. The atomic resolution capability of STM provided immediate evidence for the enormous value of this instrument as a tool for surface scientists.

STM can only be used on conductive surfaces. Given that many surfaces of technological interest are conducting or at least semiconducting, this may not seem a severe shortcoming. One might think that an STM should be capable of mapping a metallic surface under ambient conditions. However, this is not feasible because the pervasive layer of oxides and other contaminants that occurs at ambient conditions prevents stable tunneling conditions. Electrical conductivity is a necessary but not sufficient condition for a surface to be imaged by STM with atomic resolution, because the surface needs to be extremely clean on an atomic level. Except for a few extremely inert surfaces such as graphite, atomic resolution is only possible in an ultrahigh vacuum with a pressure on the order of 10^{-8} Pa and special surface preparation.

The invention of the AFM by Binnig[1], and its introduction by Binnig, Quate, and Gerber[2], opened up the possibility of obtaining true atomic resolution on conductors and insulators. Indeed, it took only a short time after the AFM's invention before apparent atomic resolution on conductors[10] and insulators[11-13] was obtained. While these early results reproduced the periodic lattice spacings of the studied samples, single defects or step edges were not observed. Also, the forces that acted between tip and sample were often orders of magnitudes larger than the forces that a tip with a single front atom was expected to be able to sustain. It was commonly assumed, therefore, that many tip atoms interacted with the surface at the same time in these early experiments. The difference between apparent and true atomic resolution of a tip with many atomic contacts can be illustrated by a macroscopic example. When profiling an egg crate with a single egg, its trajectory would represent the overall periodicity of the crate as well as each dented hump or a hole. However, when profiling one egg crate with another egg crate, again its periodicity would be retained but holes or dented humps would pass undetected. A similar effect can occur when an AFM tip probes a surface. As long as single defects, steps, or other singularities are not observed, clear proof for true atomic resolution is not established.

Even though atomic resolution was hardly ever achieved in the initial AFM experiments, the technique was readily accepted and found many technological and scientific applications. The installed base of AFMs rapidly outnumbered their STM counterparts. A recent survey[14] of the ten most highly cited publications in *Phys. Rev. Lett.* ranks the original AFM publication[2] at number four (4251 citations as of March 11, 2005 according to ISI) – in good company with

other breakthroughs in theoretical and experimental physics that have shaped our scientific life. Most of these citations refer to AFM where the spatial resolution is 'only' in the nanometer range, but the large number shows the vast range of AFM applications. In spite of the rapid growth of AFM usage, matching and even exceeding the spatial resolution of its parent, STM, had to wait for new developments.

Challenges in achieving atomic resolution

The technological foundations for achieving STM with atomic resolution (theory of electron tunneling, mechanical actuation with picometer precision, vacuum technology, surface and tip preparation, vibration isolation, etc.) were probably available a few decades before 1981, but it took the bold approach by Binnig, Rohrer, Gerber, and Weibel to pursue atomic resolution in real space. Binnig and Rohrer were rewarded with the 1986 Nobel Prize in Physics (together with Ernst Ruska, inventor of electron microscopy).

The challenges in gaining AFM with true atomic resolution are even more daunting than the hurdles that troubled STM. Fig. 1a shows a schematic view of a sharp tip for STM or AFM close to a crystalline sample, and Fig. 1b is a plot of the tunneling current and forces between tip and sample. When tip and sample are conductive and a bias voltage is applied between them, a tunneling current can flow. The red curve in Fig. 1b shows the distance dependence of the tunneling current I_t. The exponential decay of I_t with increasing distance at a rate of approximately one order of magnitude per 100 pm is the key physical characteristic that makes atomic-resolution STM possible. Because of its strong decay rate, the tunneling current is spatially confined to the front atom of the tip and flows mainly to the sample atom next to it (indicated by red circles in Fig. 1a). A second helpful property of the tunneling current is its monotonic distance

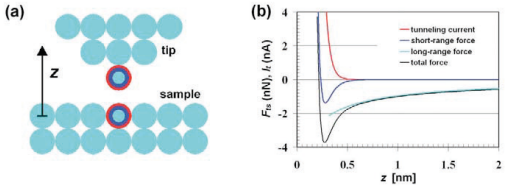


Fig. 1 (a) Schematic of tip and sample in STM or AFM. The diameter of a metal atom is typically 0.3 nm. (b) Qualitative distance dependence of tunneling current, long-, and short-range forces. Tunneling current increases monotonically with decreasing distance, while total force reaches a minimum and increases for distances below the bond length.

dependence. It is easy to build a feedback mechanism that keeps the tip at a constant distance: if the actual tunneling current is larger than the setpoint, the feedback needs to withdraw the tip and vice versa.

In contrast, the tip-sample force F_{ts} does not share the helpful characteristics of the tunneling current. First, F_{ts} is composed of long-range background forces, depicted in light-blue in Fig. 1b and originating from the atoms colored light-blue in Fig. 1a, as well as a short-range component, depicted in blue in Fig. 1b and confined to the atoms printed in blue in Fig. 1a. Because the short-range force is not monotonic, it is difficult to design a feedback loop that controls distance by using the force. A central task to perfect AFM is, therefore, the isolation of the front atom's force contribution and the creation of a linear feedback signal from it.

Even if it was possible to isolate the short-range force, a more basic problem needs to be solved first: how to measure small forces. For example, commonly known force meters, such as precise scales, are delicate and expensive instruments and even top models rarely exceed a mass resolution of 100 µg, corresponding to a force resolution of 1 µN. In addition, high-precision scales take ~1 s to acquire a weight measurement, so the bandwidth is only 1 Hz. The force meters in AFM, in contrast, require a force resolution of at least 1 nN at a typical bandwidth of 1 kHz.

Most force meters determine the deflection q' of a spring with given spring constant k that is subject to a force F with $F = q'/k$. Measuring small spring deflections is subject to thermal drift and other noise factors, resulting in a finite deflection measurement accuracy $\delta q'$. The force resolution is thus given by $\delta F = \delta q'/k$, and soft cantilevers provide less noise in the force measurement.

In contact-mode AFM, where the tip feels small repulsive forces from the sample surface, the cantilever should be softer than the bonds between surface atoms (estimated at ~10 N/m) otherwise the sample deforms more than the cantilever[15]. Because of noise and stability considerations, spring constants below 1 N/m or so have been chosen for AFM in contact mode. However, atomic forces are usually attractive in the distance regime that is best suited for atomic-resolution imaging (approximately a few hundred picometers before making contact), and soft cantilevers suffer from a 'jump-to-contact' phenomenon, i.e. when approaching the surface, the cantilever snaps toward the surface ending in an uncontrolled landing. While true atomic

resolution by contact-mode AFM has been demonstrated on chemically inert samples[16,17], this method is not feasible for imaging chemically reactive surfaces where strong, attractive short-range forces act. Long-range attractive forces are compensated in these experiments by pulling at the cantilever (negative loading force) after jump-to-contact[16] or by immersing cantilever and sample in water to reduce the van der Waals attraction[17]. Howald *et al.*[18] partially solved the reactivity problem by passivating the reactive Si tip with a thin layer of polytetrafluoroethylene (teflon). The unit cell of Si(111)-(7×7) was resolved, but atomic resolution was not reported with this method of tip passivation.

In summary, AFM shares challenges already known in STM and uses many of its design features (actuators, vibration isolation, etc.), but nature has posed four extra problems for atomic-resolution AFM: (i) jump-to-contact; (ii) nonmonotonic short-range forces; (iii) strong, long-range background forces; and (iv) instrumental noise in force measurements.

Frequency-modulation AFM

Dynamic AFM modes[19-21] help to alleviate two of the four major AFM challenges. Jump-to-contact can be prevented by oscillating the cantilever at a large enough amplitude A such that the withdrawing force on the cantilever given by $k \times A$ is larger than the maximal attractive force[22]. Because the noise in cantilever deflection measurements has a component that varies in intensity inversely with frequency ($1/f$ noise), dynamic AFM modes are less subject to noise than quasistatic operating modes. Nonmonotonic interactions and strong, long-range contributions are still present.

In amplitude-modulation AFM[19], the cantilever is driven at a constant frequency and the vibration amplitude is a measure of the tip-sample interaction. In 1991, Albrecht *et al.*[20] showed that FM-AFM offers even less noise at larger bandwidth than amplitude-modulation AFM. In FM-AFM, a cantilever with a high quality (Q) factor is driven to oscillate at its eigenfrequency by positive feedback with an electronic circuit that keeps the amplitude A constant.

A cantilever with a stiffness k and effective mass m has an eigenfrequency given by $f_0 = 1/(2\pi)\ (k/m)^{1/2}$. When the cantilever is exposed to a tip-sample force gradient k_{ts}, its frequency changes instantly to $f = f_0 + \Delta f = 1/(2\pi)\ (k'/m)^{1/2}$, where $k' = k + k_{ts}$ (Fig. 2). When k_{ts} is small compared to k, the square root can be expanded and the frequency shift is

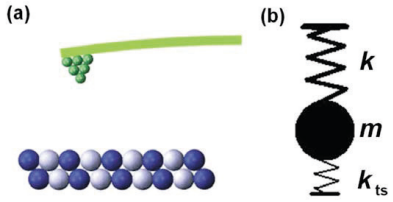


Fig. 2 (a) Schematic of a vibrating tip close to a sample in a dynamic AFM. The tip-sample forces F_{ts} cause a detectable change in the oscillation properties of the cantilever. (b) Mechanical equivalent of (a). The free cantilever with stiffness k and effective mass m has an eigenfrequency $f_0 = (k/m)^{1/2}/2\pi$. The bond between tip and sample with stiffness k_{ts} alters the resonance frequency to $f = ([k + k_{ts}]/m)^{1/2}/2\pi$. When the oscillation amplitude of the cantilever is large, k_{ts} can vary significantly within one oscillation cycle, requiring averaging.

simply given by[20]:

$$\Delta f(z) = \frac{f_0}{2k} k_{ts}(z). \qquad (1)$$

This formula is only correct if k_{ts} is constant over the distance range from $z - A$ to $z + A$ that is covered by the oscillating cantilever.

The force gradient k_{ts} was probably almost constant within the oscillation interval in the first application of FM-AFM in magnetic force microscopy by Albrecht *et al.*[20], where recording media with magnetic transitions spaced by ~2 μm were imaged using a cantilever with a stiffness of ~10 N/m oscillating at an amplitude of ~5 nm. In contrast, in the more recent application of FM-AFM in atomic-resolution AFM, k_{ts} varies by orders of magnitude throughout the oscillation of the cantilever. Using FM-AFM, true atomic resolution on Si(111)-(7×7), a fairly reactive sample, was achieved in 1994[23]. Fig. 3 shows the topographic image of this data, where the fast-scanning direction is horizontal. The atomic contrast is rather poor in the lower section, quite good in a

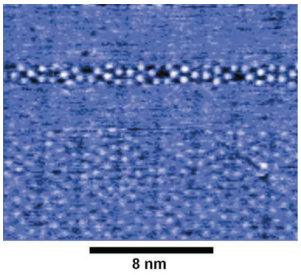


Fig. 3 First AFM image of a reactive surface showing true atomic resolution: Si(111)-(7×7) reconstruction. Parameters: k = 17 N/m; A = 34 nm; f_0 = 114 kHz; Δf = -70 Hz; Q = 28 000; and scanning speed = 3.2 lines/s. Environment: ultrahigh vacuum, room temperature. (Reprinted with permission from[23]. © 1995 AAAS.)

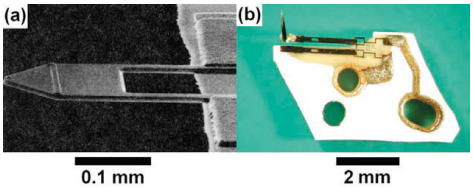


Fig. 4 Micrographs of (a) a piezoresistive cantilever[24] and (b) a 'qPlus' sensor[46] – a cantilever made from a quartz tuning fork. The piezoresistive cantilever is 250 μm long, 50 μm wide, and 4 μm thick. The eigenfrequency is 114 kHz, the stiffness 17 N/m, and the Q factor in vacuum 28 000. The qPlus sensor has a typical eigenfrequency ranging from 10-30 kHz (depending on the mass of the tip), a stiffness of 1800 N/m, and a Q factor of 4000 in vacuum at T = 300K and 20 000 at T = 4K. One of the prongs is fixed to a large substrate and a tip is mounted to the free prong. Because the fixed prong is attached to a heavy mass, the device is mechanically equivalent to a traditional cantilever. The free prong is 2.4 mm long, 130 μm wide, and 214 μm thick.

narrow strip in the center, and vanishing in the top section. These changes in contrast are the result of tip changes, indicating fairly strong interaction during the imaging process. A piezoresistive cantilever made of Si[24], as shown in Fig. 4a, with a stiffness of 17 N/m was used to obtain this image. The amplitude of the cantilever can be freely adjusted by the operator and, while it was planned to use the thermally excited amplitude[25] (~10 pm), the empirically determined optimal amplitude values were always around 10 nm – a similar order of magnitude as the value of A = 34 nm used in Fig. 3. The chemical bonding forces that are responsible for the atomic contrast in AFM imaging of Si have a range on the order of 100 pm[26], so the amplitude is 340 times as large.

The requirement of such a large amplitude is in stark contrast to intuition. Imagine an atom magnified to a size of an orange with a diameter of 8 cm. The range of the bonding force is then only 4 cm or so. The front atom of the cantilever approaches from a distance of 20 m and only in the last few centimeters of its oscillation cycle does it feel the attractive bonding forces from the sample atom next to it. On the other hand, force gradients can be quite large in chemical bonds. According to the well-known Stillinger-Weber potential[27], a single bond between two Si atoms has a force gradient of $k_{ts} \approx +170$ N/m at the equilibrium distance of z = 235 pm and $k_{ts} \approx -120$ N/m when the two Si atoms are at a distance of z = 335 pm. Because of the relatively large values of interatomic force gradients, even cantilevers with a stiffness on the order of 1 kN/m should be subject to significant frequency shifts when oscillating at small amplitudes[28]. Nevertheless, the large-amplitude FM-AFM technique has celebrated great successes in imaging metals, semiconductors, and insulators with true atomic resolution[29-33].

Optimal imaging parameters

In order to understand why these large oscillation amplitudes are necessary, a quantitative analysis of the physics of large-amplitude FM-AFM is necessary, starting with a calculation of frequency shift for large amplitudes. If k_{ts} is not constant over one oscillation cycle, eq 1 no longer holds and a perturbation theory, such as the Hamilton-Jacobi theory[34], can be used to find the relationship between frequency and tip-sample forces[22]. Other perturbative approaches have confirmed the result[35-38], and an instructive representation of the formula is:

$$\Delta f(z) = \frac{f_0}{\pi k} \int_{-1}^{1} k_{ts}(z - uA)\sqrt{1-u^2}\,du. \qquad (2)$$

This equation is key to a physical understanding of FM-AFM, allowing evaluation of the impact of various force components on Δf, the experimental observable. At first glance, the large-amplitude result resembles eq 1, where $k_{ts}(z)$ is replaced by an averaged value. The average force gradient is computed by convoluting $k_{ts}(z)$ in the interval $z - A$ to $z + A$ with a semispherical weight function. The weight function has its maximum at u = 0, a distance A away from the minimal tip-sample distance. The minimal tip-sample distance z_{min} is an important parameter in any STM or AFM experiment because, while a small value of z_{min} is desirable for optimal spatial resolution, both tip and sample can be damaged if z_{min} is too small. We can now ask, if we keep z_{min} constant and vary A, what happens to our signal, the frequency shift Δf? The answer is given in eq 2: as long as the gradient of the tip-sample interaction k_{ts} remains constant as the tip of the cantilever moves over a z-range from z_{min} to $z_{min} + 2A$, Δf stays constant. However, as A reaches the decay length λ of the interaction, the frequency shift drops sharply at a rate $\propto (\lambda/A)^{3/2}$. It turns out[39] that, for amplitudes larger than λ, Δf is no longer proportional to the force gradient, but to the product of force and the square root of λ (or, equivalently, to the geometric average between potential and force[40]). In FM-AFM with amplitudes large compared to the interaction range, it is useful to define a quantity[22] $\gamma = \Delta f k A^{3/2}/f_0$. The 'normalized frequency shift' γ connects the physical observable Δf and the underlying forces F_{ts} with range λ, where $\gamma \approx 0.4 F_{ts}\lambda^{1/2}$ (see eqs 35-41 in[33]). For covalent bonds, the typical bonding strength is on the

order of -1 nN with $\lambda \approx 1$ Å, resulting in $\gamma \approx -4$ fNm$^{1/2}$, where a negative sign indicates attractive interaction. The crossover from the small-amplitude approximation in eq 1 to the large-amplitude case in eq 2 occurs for amplitudes on the order of the interaction range λ.

Eq 2 determines the influence of the oscillation amplitude on the third AFM challenge: the disturbing contribution of long-range forces. Imagine an AFM tip at a minimal distance z_{min} = 0.3 nm from a surface, where the total tip-sample force is composed of a chemical bonding force with an exponential distance dependence and a given range, plus a long-range force with the same strength and a ten times longer range (see Table 1 for details). In large-amplitude AFM (here, A > 1 nm), the signal is proportional to γ, and the long-range contribution to Δf is (1 nm/100 pm)$^{1/2}$, or approximately three times larger than the short-range contribution. For small amplitudes (here, A < 100 pm), Δf is proportional to the force gradient and the long-range component is only 100 pm/1 nm, or 1/10 of the short-range contribution. Therefore, small-amplitude AFM helps to reduce the unwanted contribution of long-range forces.

Even stronger attenuation of the unwanted long-range contribution would be possible if higher-order force derivatives could be mapped directly. For example, if we could directly measure $\partial^2 F_{ts}/\partial z^2$, the long-range component would be only 1/100 of the short-range contribution. For a direct mapping of the third-order gradient $\partial^3 F_{ts}/\partial z^3$, the relative long-range component would reduce to a mere 1/1000. Higher force gradients can be mapped directly by higher harmonic AFM, as described below.

Because the forces that act in AFM are small, optimizing the signal-to-noise ratio is crucial for obtaining good images. Frequency noise in FM-AFM is inversely proportional to amplitude[19,20,33,41]. As discussed above, the signal stays constant until A reaches λ and drops proportional to $(\lambda/A)^{3/2}$ for larger amplitudes. Therefore, the signal-to-noise ratio is maximal for amplitudes on the order of the decay length of the interaction that is used for imaging[42]. For atomic imaging, amplitudes on the order of 100 pm are expected to be optimal. As a conclusion of these calculations, we find that the use of small amplitudes $A \approx \lambda$ would have two advantages: (i) increased signal-to-noise ratio[42]; and (ii) greater sensitivity to short-range forces[33].

So, why was it not feasible to use small amplitudes in the initial experiments? Two reasons, related to the mechanical stability of the oscillating cantilever, can be identified. First, jump-to-contact is prevented if the withdrawing force of the cantilever when it is closest to the sample given by $k \times A$ is larger than the maximal attraction[22]. Second, because tip-sample forces are not conservative[43], random dissipative phenomena with a magnitude of δE_{ts} cause amplitude fluctuations[42,44] $\delta A = \delta E_{ts}/(kA)$. Both problems can be resolved by using cantilevers with sufficient stiffness. Stability considerations propose a lower threshold for k that depends on the tip-sample dissipation as well as the Q factor of the cantilever. Because the frequency shift is inversely proportional to the stiffness (eqs 1 and 2), k should still be as low as permitted by the stability requirements.

Stiff cantilevers were not available when we realized their potential advantages; therefore, we built cantilevers with a stiffness of k = 1800 N/m from quartz tuning forks[44-46] (Fig. 4b). A secondary advantage of quartz cantilevers is their greater frequency stability with temperature, which leads to lower frequency drift, particularly if a quartz-stabilized frequency detector is used (we used the EasyPLL by Nanosurf®, Switzerland). Other small-amplitude approaches with stiff, home-built W cantilevers have been demonstrated by the Erlandsson[47] and Pethica[48-50] groups.

As predicted by theoretical considerations, the stiff cantilever allows use of subnanometer amplitudes, resulting

Table 1 *Short- and long-range contributions to AFM signals in different operating modes.*

AFM method	Physical observable	Short-range contribution	Long-range contribution	Relative short-range contribution
Quasistatic	force	1 nN	1 nN	50%
Large-amplitude FM	$\gamma \approx 0.4 \times \text{force} \times \sqrt{\text{range}}$	4 fNm$^{1/2}$	12 fNm$^{1/2}$	25%
Small-amplitude FM	force gradient	10 N/m	1 N/m	91%
Higher-harmonic	n^{th} force gradient	$10^{n+9(n-1)}$ N/m^n	$10^{9(n-1)}$ N/m^n	$\approx 100\%(1\text{-}10^{-n})$

This model calculation assumes a chemical bonding force $F(z) = F_0 e^{-z/\lambda}$ with a strength $F_{short\ range}(z_{min}) = 1$ nN and range $\lambda_{short\ range} = 100$ pm, as well as an equally strong long-range background force with $F_{long\ range}(z_{min}) = 1$ nN and a range of $\lambda_{long\ range} = 1$ nm. Depending on the mode of AFM operation, the short-range part has a different weight in the total interaction signal. Higher-harmonic AFM offers the greatest attenuation of long-range forces.

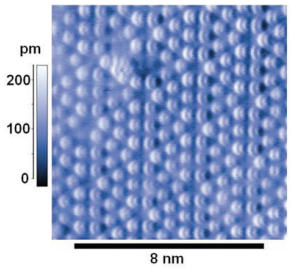


Fig. 5 AFM image of the Si 7×7 reconstruction with true atomic resolution using a stiff cantilever. Parameters: k = 1800 N/m; A = 0.8 nm; f_0 = 16.86 kHz; Δf = -160 Hz; and Q = 4000. Environment: ultrahigh vacuum, room temperature. (Reprinted with permission from[51]. © 2000 AAAS.)

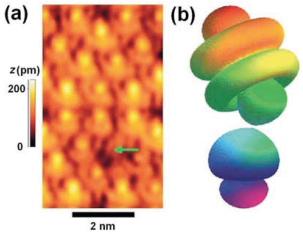


Fig. 6 (a) Dynamic STM image of the Si 7×7 reconstruction using a Co_6Fe_3Sm tip mounted on a qPlus sensor. Parameters: k = 1800 N/m; A = 0.5 nm, f_0 = 19 621 Hz; sample bias voltage = -100 mV; and average tunneling current = 200 pA. Environment: ultrahigh vacuum, room temperature. (b) Schematic of tip and sample states that can lead to the experimental image shown in (a). The sample state is a dangling bond of a Si adatom with $3sp^3$ symmetry, while a Sm $4f_{z_3}$ state is taken as the tip state. (Reprinted with permission from[55,56]. © 2003 American Physical Society.)

in an improved signal-to-noise ratio, strong attenuation of long-range forces, and stable scanning at very small tip-sample distances. For these reasons, spatial resolution is increased, as shown in Fig. 5. The image shows a very clear picture of Si with a defect and very large corrugation. The adatoms of Si, which should be spherically symmetric, show subatomic details that are interpreted as orbitals in the tip atom[51,52]. This AFM image seems to show greater resolution than what was known from STM. According to the 'Stoll formula'[53], a theoretical estimate of the vertical corrugation and, thus, the lateral resolution of STM images, two physical parameters are crucial for the high spatial resolution of STM: (i) the very short decay length of the tunneling current; and (ii) a small tip-sample distance. Three likely reasons have been identified that may explain why dynamic AFM might provide better resolution than STM[54]:

- In dynamic AFM, the minimal tip-sample distance can be much smaller than in STM without destroying the tip because the shear forces that act on the front atom during scanning are much smaller in the oscillation phase where the tip is far from the sample.
- When using large gap voltages, a variety of states can contribute to the tunneling current, smearing the image.
- Tip-sample forces also have repulsive components with a very short decay length.

The first two characteristics can also be fulfilled in STM by using a very small tunneling bias voltage and oscillating the STM tip. Fig. 6 shows an image of Si obtained using dynamic STM, where a Co_6Fe_3Sm magnetic tip was mounted onto a qPlus sensor[55,56]. Each Si adatom looks like a fried egg with a sharp central peak surrounded by a halo. The radius of the central peak is only on the order of 100 pm, showing that higher-momentum states[57] must have been involved in this image. The experiment was repeated with pure Co, Fe, and Sm tips, and only pure Sm tips yielded similar images to Fig. 6. We conclude, therefore, that a Sm atom acts as the tip atom in this experiment[55]. In atomic Sm, the electrons in the highest occupied state are in a *4f* state. If one assumes that the electronic states at a Sm surface atom of bulk Co_6Fe_3Sm are similar to atomic states in Sm, it appears likely that the crystal field around the front atom creates a state close to $4f_{z_3}$ symmetry that is responsible for the tunneling contrast. Interestingly, very small tip-sample distances can only be realized with oscillating tips. When the oscillation is turned off, the current setpoint has to be reduced, otherwise the tip would not survive the small tunneling distances.

Operation at small oscillation amplitudes not only results in greater resolution, it also facilitates simultaneous STM and AFM imaging. A straightforward implementation of combined current and force measurements uses the constant-height mode, where the *z*-position of the tip is held constant relative to the plane connecting the surface atoms. A simultaneous measurement of tunneling current and frequency shift allows comparison of the forces and tunneling currents. Fig. 7 shows the current and repulsive force on graphite[58] observed by simultaneous AFM and STM in vacuum at liquid helium temperatures (4.9 K). STM only sees the electrons at the Fermi level, while repulsive forces act wherever the local charge density is high (i.e. over every atom) for small enough distances. In graphite, only every second surface atom conducts electricity, but every surface atom exerts repulsive forces. Therefore, AFM 'sees more' than STM and allows correlation of topography with local

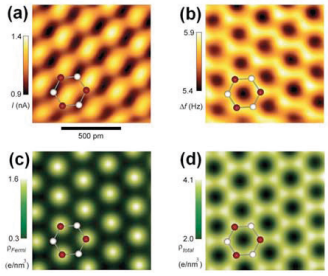


Fig. 7 (a) Constant-height STM image of graphite, and (b) simultaneously recorded AFM image (repulsive). (c) Estimate of the charge density at the Fermi level (visible in STM), and (d) total charge density (relevant for repulsive AFM) for graphite. Parameters: k = 1800 N/m; A = 0.3 nm; f_0 = 18 076.5 Hz; and Q = 20 000. (Reprinted with permission from[58]. © 2003 National Academy of Sciences, USA.)

conductance. This method is promising for other materials with more than one basis atom in the elementary cell[59,60].

While a strong bias dependence holds for atomic-resolution STM[61] as well as AFM images[60,62], a pronounced difference is that the tunneling current direction is not accessible in STM, while the measured force direction is determined by the cantilever's orientation. Usually, AFM senses forces normal to the surface, but it is also possible to perform lateral force microscopy[63] by measuring forces parallel to the surface. In a quasistatic mode, lateral forces can be recorded simultaneously with normal forces. In dynamic modes, it is easier to rotate the attachment of the cantilever by 90° and detect lateral forces. Fig. 8 shows a measurement of lateral force gradients between a tip and a Si surface. Parallel motion between tip and cantilever also allows the use of extremely soft cantilevers to probe the limits of force resolution without suffering jump-to-contact, as shown by Rugar *et al.*[64] in single-spin detection by magnetic resonance force microscopy.

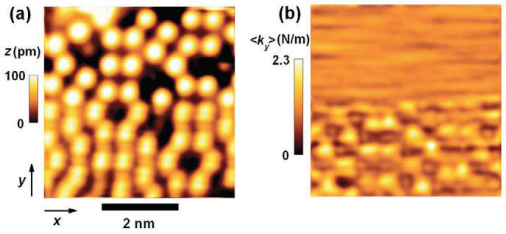


Fig. 8 (a) Topographic STM image of Si(111)-(7×7) where the tip is mounted on a lateral force sensor. The tip oscillates with A ≈ 80 pm in the y-direction in the lower half of the image; the oscillation is turned off in the upper half. (b) Corresponding lateral force gradient. On top of the adatoms, the bond between tip and sample causes an increase in frequency shift. Parameters: k = 1350 N/m; A = 80 pm (bottom), A = 0 (top); f_0 = 10 214 Hz. Environment: ultrahigh vacuum, room temperature[56].

Higher-harmonic AFM

Can we increase the spatial resolution of AFM any further? When decreasing the amplitude from A >> λ to A << λ, the frequency shift changes from a proportionality of $F_{ts}\lambda^{1/2}$ to F_{ts}/λ. As outlined above, an experimental observable that is proportional to a higher force gradient should allow even higher spatial resolution than small-amplitude FM-AFM. Luckily, there is a physical observable that couples directly to higher force gradients. When the cantilever oscillates in the force field of the sample, a shift in frequency is not the only change in the cantilever's motions. The oscillation of the cantilever changes from a purely sinusoidal motion, given by $q' = A\cos(2\pi ft)$, to an oscillation that contains higher harmonics with $q' = \Sigma_{n=0}^{\infty} a_n\cos(2\pi nft + \phi_n)$. For amplitudes that are large with respect to the range of F_{ts}, the higher harmonics are essentially proportional[37] to Δf. However, for small amplitudes, Dürig[65] has found that F_{ts} can be recovered immediately within the distance range from z_{min} to $z_{min} + 2A$ if the amplitudes and phases of all higher harmonics of the cantilever's motion are known. Moreover, higher harmonics bear even more useful information: direct coupling to higher force gradients[66]. Similar to eq 2, we can express the magnitude of the higher harmonics by a weighted average of a force gradient – a gradient of order $n > 1$ this time:

$$a_n = \frac{2}{\pi k}\frac{1}{1-n^2}\frac{A^n}{1\cdot 3\cdot \ldots \cdot (2n-1)}\int_{-1}^{1}\frac{d^n F_{ts}(z+Au)}{dz^n}(1-u^2)^{n-1/2}du. \quad (3)$$

The weight function changes from the semispherical shape $w_{\Delta f}(u) = (1 - u^2)^{1/2}$ in eq 2 to functions $w_n(u) = (1 - u^2)^{n-1/2}$ that are more and more peaked with increasing n. For this reason, the use of small amplitudes is of even greater importance in higher harmonic AFM than in FM-AFM. The magnitude of the higher harmonic amplitudes a_n is rather small compared to the fundamental amplitude $a_1 = A$; therefore, higher harmonic AFM works best at low temperatures, where the detection bandwidth can be set to very small values.

The spatial resolution of AFM and STM is fundamentally limited neither by the mechanical vibration level nor by thermal vibrations, but by the spatial extent of the experimental objects that are observed – electrons at the Fermi level in STM[67] and something close to the total charge density in repulsive AFM[68]. When probing the resolution limits of AFM, we first have to find an object with the desired sharply localized electronic states. Pauling[69] noted that

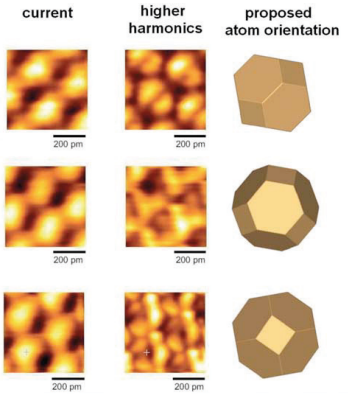


Fig. 9 Simultaneous constant-height STM (left column) and higher-harmonic AFM images (central column) of graphite using a W tip. The right column shows the proposed orientation of the W tip atom. The W atom is represented by its Wigner-Seitz unit cell, which reflects the full symmetry of the bulk. We assume that the bonding symmetry of the adatom is similar to the bonding symmetry of the bulk. This assumption is based on charge density calculations of surface atoms[70,71]. In the first row, the higher harmonics show a two-fold symmetry, resulting from a [110] orientation of the front atom. In the second row, the higher harmonics show a roughly three-fold symmetry, as expected for a [111] orientation. In the third row, the symmetry of the higher-harmonic signal is approximately four-fold, as expected for a tip in [001] orientation. Parameters: k = 1800 N/m; A = 0.3 nm; f_0 = 18 076.5 Hz; and Q = 20 000. Environment: ultrahigh vacuum, T = 4.9 K. (Reprinted with permission from[66]. © 2004 AAAS.)

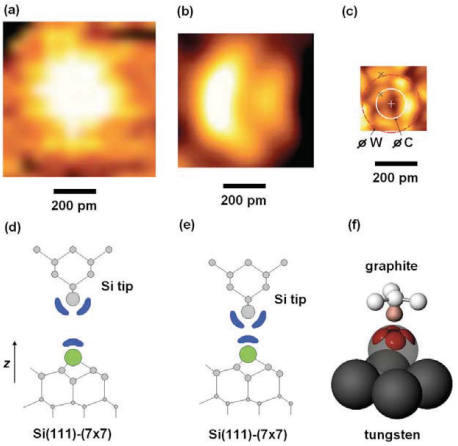


Fig. 10 Progress in AFM spatial resolution showing images of single atoms. The lateral scale in (a)-(c) is equal. (a) An adatom of the Si(111)-(7×7) reconstruction, showing up as a blurred spot. (b) An adatom of the Si(111)-(7×7) reconstruction, showing subatomic contrast originating in the electronic structure of the tip. (c) Higher-harmonic image of a W atom mapped by a carbon atom. Parameters: (a) k = 17 N/m; A = 34 nm; f_0 = 114 kHz, Δf = -70 Hz; and Q = 28 000 (ultrahigh vacuum, room temperature); (b) k = 1800 N/m; A = 0.8 nm; f_0 = 16 860Hz; Δf = -160 Hz; and Q = 4 000 (ultrahigh vacuum, room temperature); (c) k = 1800 N/m; A = 0.3 nm; f_0 = 18 076.5 Hz; and Q = 20 000 (ultrahigh vacuum, T = 4.9 K), higher harmonic detection. (d) Schematic of a Si(001) tip close to a Si(111)-(7×7) surface. Because of the large amplitude and a fairly large minimum tip-sample distance, the blurry image (a) corresponding to this configuration is approximately symmetric with respect to the vertical axis. (e) Similar to (d), but at a closer distance. The angular dependence of the bonding forces is noticeable. (f) W(001) surface close to a carbon atom in a graphite surface. The charge distribution in W shows small pockets that are resolved by higher-harmonic AFM with a light-atom carbon-probe. (Parts (c) and (f) reprinted with permission from[66]. © 2004 AAAS.)

transition metals show a covalent bonding character and should therefore expose lobes of increased charge density toward their neighbors. Indeed, while the surface atoms of W(001) expose a large blurred charge cloud at the Fermi level for *k*-vectors perpendicular to the surface (Fig. 8 in[70]), the total charge density shows four distinct maxima (Fig. 3 in[70] and Fig. 3a in[71]). Fig. 9 shows a direct comparison of the simultaneously recorded tunneling current and higher harmonic amplitudes. As expected, the higher harmonic data shows much greater detail.

Summary and conclusion

We have substantiated the enormous usefulness of AFM by referring to the numerous references to the original publication[2] in the introduction. While most AFM applications are currently not in the atomic-resolution regime, the enhancement in spatial resolution is likely to create value in most AFM studies in physics, chemistry, biology, and materials science.

Recently, true atomic resolution in FM-AFM has been observed at ambient pressure in an N_2 atmosphere[72], showing that some of the concepts of vacuum AFM are applicable in ambient environments. Although STM resolution can benefit from oscillating the tip, a concept that originated in AFM, Fig. 9 shows that AFM has now clearly reached and even surpassed the resolution capability of STM. Fig. 10 shows the evolution of AFM resolution from large-amplitude AFM in 1994 (Fig. 10a) to small-amplitude AFM in 2000 (Fig. 10b) and higher-harmonic AFM in 2004 (Fig. 10c). While the structures within single atoms shown in Figs. 10b and 10c originate in the front atom of the probe, there are other examples where AFM shows more atomic details than STM that establish the improved spatial resolution of AFM over STM in special cases. These include the observation of the rest atoms in Si(111)-(7×7)[73,74] or the observation of all dangling bonds on the Si/Ge(105) surface[75].

Atomic and molecular structuring has long been the domain of STM, from the first demonstration of manipulating single atoms[76] to a variety of STM nanofabrication methods[77]. Recently, it has been shown that atomic manipulation by AFM is possible even at room temperature[78].

We have not been able to discuss the phenomenal success of AFM in biology, a field with a much more immediate impact on the human condition. It can be expected that at least some of the concepts that have been developed for AFM in vacuum will enable greater resolution in biological AFM applications as well[79,80]. MT

Acknowledgments

I wish to thank Jochen Mannhart for support and editorial suggestions and current and past students Martin Breitschaft, Philipp Feldpausch, Stefan Hembacher, Markus Herz, Christian Schiller, Ulrich Mair, Thomas Ottenthal, and Martina Schmid, as well as lab engineers Klaus Wiedenmann and Alexander Herrnberger, for contributing to the progress of AFM. I also thank Gerd Binnig, Calvin F. Quate, and Christoph Gerber for kicking off the fun of AFM and for ongoing inspiring interactions. Special thanks to Heinrich Rohrer, Calvin Quate, and Christoph Gerber for critical comments, and German Hammerl for help with LaTeX. Supported by the Bundesministerium für Forschung und Technologie (contract 13N6918).

REFERENCES

1. Binnig, G., Atomic Force Microscope and Method for Imaging Surfaces with Atomic Resolution, US Patent 4,724,318, (1986)
2. Binnig, G., *et al., Phys. Rev. Lett.* (1986) **56** (9), 930
3. Binnig, G., and Rohrer, H., *Rev. Mod. Phys.* (1999) **71** (2), S324
4. Faraday, M., *Chemical History of a Candle*, Dover Publications, Mineola, NY, (2003)
5. Siegman, A. E., *Lasers*, University Science Books, Herndon, VA, (1986)
6. Binnig, G., *et al., Phys. Rev. Lett.* (1982) **49** (1), 57
7. Hofer, W. A., *Materials Today* (2002) **5** (10), 24
8. Binnig, G., *et al., Phys. Rev. Lett.* (1983) **50** (2), 120
9. Takayanagi, K., *et al., J. Vac. Sci. Technol. A* (1985) **3** (3), 1502
10. Binnig, G., *et al., Europhys. Lett.* (1987) **3**, 1281
11. Albrecht, T. R., and Quate, C. F., *J. Appl. Phys.* (1987) **62** (7), 2599
12. Meyer, G., and Amer, N. M., *Appl. Phys. Lett.* (1990) **56** (21), 2100
13. Meyer, E., *et al., Z. Phys. B: Condens. Matter* (1990) **79**, 3
14. Riordon, J., *APS News* (2003) May, 3
15. Rugar, D., and Hansma, P., *Phys. Today* (1990) **43**, 23
16. Giessibl, F. J., and Binnig, G., *Ultramicroscopy* (1992) **42-44** (1), 281
17. Ohnesorge, F., and Binnig, G., *Science* (1993) **260**, 1451
18. Howald, L., *et al., Phys. Rev. B* (1995) **51** (8), 5484
19. Martin, Y., *et al., J. Appl. Phys.* (1987) **61** (10), 4723
20. Albrecht, T. R., *et al., J. Appl. Phys.* (1991) **69** (2), 668
21. Dürig, U., *et al., J. Appl. Phys.* (1992) **72** (5), 1778
22. Giessibl, F. J., *Phys. Rev. B* (1997) **56** (24), 16010
23. Giessibl, F. J., *Science* (1995) **267**, 68
24. Tortonese, M., *et al., Appl. Phys. Lett.* (1993) **62**, 834
25. Giessibl, F. J., *Jpn. J. Appl. Phys., Part 1* (1994) **33** (6B), 3726
26. Perez, R., *et al., Phys. Rev. Lett.* (1997) **78** (4), 678
27. Stillinger, F. H., and Weber, T. A., *Phys. Rev. B* (1985) **31** (8), 5262
28. Giessibl, F. J., German Patentschrift DE 196 33 546, (1996)
29. Morita, S., *et al.* (eds.), *Noncontact Atomic Force Microscopy*, Springer, Berlin, (2002)
30. Meyer, E., *et al., Scanning Probe Microscopy: The Lab on a Tip*, Springer, Berlin, (2003)
31. Garcia, R., and Perez, R., *Surf. Sci. Rep.* (2002) **47** (6-8), 197
32. Hofer, W. A., *et al., Rev. Mod. Phys.* (2003) **75**, 1287
33. Giessibl, F. J., *Rev. Mod. Phys.* (2003) **75**, 949
34. Goldstein, H., *Classical Mechanics*, Addison Wesley, Reading, MA (1980)
35. Baratoff, A., unpublished results (1997)
36. Dürig, U., *Surf. Interface Anal.* (1999) **27** (5-6), 467
37. Dürig, U., *Appl. Phys. Lett.* (1999) **75** (3), 433
38. Livshits, A. I., *et al., Appl. Surf. Sci.* (1999) **140** (3-4), 327
39. Giessibl, F. J., and Bielefeldt, H., *Phys. Rev. B* (2000) **61** (15), 9968
40. Ke, S. H., *et al., Phys. Rev. B* (1999) **59** (20), 13267
41. Hasegawa, Y., *et al., Jpn. J. Appl. Phys.* (2004) **43** (2B), L303
42. Giessibl, F. J., *et al., Appl. Surf. Sci.* (1999) **140** (3-4), 352
43. Kantorovich, L. N., *Phys. Rev. B* (2001) **64** (24), 245409
44. Giessibl, F. J., *et al., Nanotechnology* (2004) **15** (2), S79
45. Giessibl, F. J., *Appl. Phys. Lett.* (1998) **73** (26), 3956
46. Giessibl, F. J., *Appl. Phys. Lett.* (2000) **76** (11), 1470
47. Erlandsson, R., *et al., Phys. Rev. B* (1997) **54** (12), R8309
48. Hoffmann, P. M., *et al., Proc. R. Soc. London, Ser. A* (2001) **457**, 1161
49. Hoffmann, P. M., *et al., Phys. Rev. Lett.* (2001) **87** (26), 265502
50. Oral, A., *et al., Appl. Phys. Lett.* (2001) **79** (12), 1915
51. Giessibl, F. J., *et al., Science* (2000) **289**, 422
52. Huang, M., *et al., Phys. Rev. Lett.* (2003) **90**, 256101
53. Stoll, E., *Surf. Sci. Lett.* (1984) **143** (2-3), L411
54. Giessibl, F. J., *et al., Ann. Phys. (Leipzig)* (2001) **10** (11-12), 887
55. Herz, M., *et al., Phys. Rev. B* (2003) **68**, 045301
56. Herz, M., Dynamische Tunnel-, Kraft- und Reibungsmikroskopie mit atomarer und subatomarer Auflösung, PhD thesis, University of Augsburg, Germany, (2003), www.lob.de
57. Chen, C. J., *Introduction to Scanning Tunneling Microscopy*, Oxford University Press, New York, (1993)
58. Hembacher, S., *et al., Proc. Natl. Acad. Sci. USA* (2003) **100** (22), 12539
59. Hembacher, S. F., Simultane Rasterkraft- und Rastertunnelmikroskopie bei 5 K im Ultrahochvakuum, PhD thesis, University of Augsburg, Germany, (2003), www.dissertation.de
60. Hembacher, S., *et al., Phys. Rev. Lett.* (2005) **94**, 056101
61. Feenstra, R. M., *et al., Phys. Rev. Lett.* (1987) **58** (12), 1192
62. Arai, T., and Tomitori, M., *Phys. Rev. Lett.* (2004) **93**, 256101
63. Mate, C. M., *et al., Phys. Rev. Lett.* (1987) **59** (17), 1942
64. Rugar, D., *et al., Nature* (2004) **430**, 329
65. Dürig, U., *New J. Phys.* (2000) **2**, 5
66. Hembacher, S., *et al., Science* (2004) **305**, 380
67. Tersoff, J., and Hamann, D. R., *Phys. Rev. Lett.* (1985) **50** (25), 1998
68. Ciraci, S., *et al., Phys. Rev. B* (1990) **41** (5), 2763
69. Pauling, L., *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, NY, (1957)
70. Posternak, M., *et al., Phys. Rev. B* (1980) **21** (12), 5601
71. Mattheiss, L. F., and Hamann, D. R., *Phys. Rev. B* (1984) **29** (10), 5372
72. Sasahara, A., *et al., J. Phys. Chem. B* (2004) **108** (40), 15735
73. Lantz, M. A., *et al., Phys. Rev. Lett.* (2000) **84** (12), 2642
74. Eguchi, T., and Hasegawa, Y., *Phys. Rev. Lett.* (2002) **89**, 266105
75. Eguchi, T., *et al., Phys. Rev. Lett.* (2004) **93**, 266102
76. Eigler, D. M., and Schweizer, E. K., *Nature* (1990) **344**, 524
77. Rosei, F., *J. Phys.: Condens. Matter* (2004) **16** (17), S1373
78. Sugimoto, Y., *et al., Nat. Mater.* (2005) **4** (2), 156
79. Miles, M., *Science* (1997) **277**, 1845
80. Hörber, J. K. H., and Miles, M. J., *Science* (2003) **302**, 1002

Femtosecond X-ray diffraction

Thomas Elsässer
Max-Born-Institute, Berlin, Germany
elsasser@mbi-berlin.de

Numerous nonequilibrium processes in condensed matter involve changes of atomic and electronic structure. The rearrangement of atoms in a phase transition or a chemical reaction, charge relocations and changes of spin states in electron transfer and/or magnetic processes, as well as field-driven changes and mixing of electronic orbitals are basic phenomena which determine functional properties. The elementary steps of many of such processes occur on ultrafast time scales, covering a range from 100 as (10^{-16} s) to ~10 ps (10^{-11} s). Nonlinear time-resolved spectroscopy from the far infrared to the soft x-ray range has given detailed insight into such dynamics and the interactions governing them but provides very limited information on structures at atomic length and time scales.
In recent years, there has been impressive progress in developing and applying x-ray probes of ultrafast structural dynamics. Femtosecond x-ray diffraction has seen rapid progress with numerous applications in physics, (bio)chemistry, and materials science. This lecture gives an introduction into this new field with particular emphasis on structural dynamics in crystalline materials. After introducing basic experimental concepts, prototypical results on lattice and charge dynamics are presented, with particular emphasis on time-resolved x-ray powder diffraction data for deriving time-dependent charge density maps and electric polarizations. The present lecture notes provide some background and are partly based on the review articles [1-4].

Sources of ultrashort hard x-ray pulses and experimental methods

X-ray sources. X-ray diffraction requires hard x-ray radiation with wavelengths smaller than or at least comparable to interatomic distances in matter. There are two different classes of sources of ultrashort x-ray pulses, (i) accelerator based large scale facilities such as free electron lasers (FELs) and so-called slicing beamlines at synchrotrons [5-7], and (ii) laser-driven table-top plasma sources. The present generation of hard x-ray FELs is based on the self-amplified stimulated emission (SASE) of x-rays generated by relativistic electron bunches in undulator structures [8]. X-ray generation starts with spontaneous radiation from an electron bunch (macro-bunch) which is structured into micro-bunches by interaction with the strong radiation field after a sufficient propagation length in the undulator. The micro-bunches display a spatial periodicity identical to the light wave and, thus, the coherent superposition of the emission from all micro-bunches leads to an intensity $I \propto N^2$ where N is the total number of radiating electrons. The first hard x-ray FEL has been the Linac Coherent Light Source (LCLS) at SLAC, Stanford [9], followed by SACLA, Sayo, Japan [10], the European XFEL in Hamburg, and a number of

ongoing projects, among them the SwissFEL, Villigen. Some key parameters of the LCLS hard x-ray output are summarized in Table 1.

The statistically fluctuating initial condition of the SASE process results in fluctuations of both the spectral and the time structure of the generated pulses. The time structure of the SASE FEL pulses consists of a sequence of up to several 100 short coherent spikes with an average coherence time of 0.55 fs and coherence length of 17 μm [11]. The timing jitter between the SASE FEL pulses and an external femtosecond laser is of the order of 200 fs r.m.s. (root mean square), affecting the time resolution of optical pump/x-ray probe experiments. The spectral and temporal pulse characteristics can be strongly improved by seeding the FEL, i.e., providing a well-defined initial condition for the amplification process. Hard x-ray self seeding of the LCLS has been reported in ref. [12].

Laser-driven hard x-ray sources rely on the interaction of femtosecond laser pulses with a peak intensity of the order of $I=10^{17}$ W/cm^2 with a metallic target [13]. The strong electric field of the pulse extracts electrons from the metal surface by field ionization. The generated free electrons are subsequently accelerated for a half cycle of the driving field into the vacuum, a process called vacuum heating [14], and in the next half cycle of opposite sign smashed back into the target. The electrons reach a kinetic energy of several hundreds of keV which is proportional to $I\lambda^2$ (λ : optical driving wavelength). They generate characteristic x-ray emission by inner shell ionization of target atoms and Bremsstrahlung by inelastic scattering with target atoms. The time structure of the characteristic x-ray emission is determined by the duration of the driving pulse and the electron deceleration kinetics in the target, in particular by the target thickness.

State-of-the-art hard x-ray sources are driven by sub-50 fs pulses of millijoule energies from amplified Ti:sapphire lasers working at a 1 kHz repetition rate [15,16]. Using Cu tape targets of 20 μm thickness, characteristic Kα pulses (photon energy 8.04 keV) have been generated with a total flux of up to 10^{11} photons/s into the full solid angle. For diffraction experiments, x-ray emission into a fraction of the total solid angle is collected with an x-ray optics and focused onto the sample, resulting in a collimated hard x-ray flux of up to several 10^6 photons/s. The duration of the hard x-ray pulses is 100 fs for a target thickness of 20 μm. A major advantage of this generation scheme consists in the synchronization of the x-ray pulses with the driving laser output, i.e. a timing jitter negligible compared to the x-ray pulse duration. Table 1 summarizes the main parameters of such sources.

An enhancement of the kinetic energy of electrons in the vacuum heating process results in a higher yield of both characteristic x-ray emission and Bremsstrahlung. This is possible with driver pulses of longer wavelength which provide a longer electron acceleration period. An enhancement of the x-ray photon number per pulse by a factor of 20 was observed with driving pulses centered at a wavelength of 4 μm [17], in agreement with theoretical simulations [18]. The development of mid-infrared drivers with kilohertz repetition rates is presently underway [19].

TABLE I. Femtosecond x-ray sources. Pulse parameters are summarized for the hard x-ray output of LCLS at SLAC, the slicing beamline at the Swiss Light Source (SLS), and laser-driven Cu and Mo Kα plasma sources at the Max Born Institute (MBI) in Berlin.

	FEL (LCLS)	Slicing (SLS)	Laser plasma (MBI)
Photon energy E (keV)	2.0–9.6	4.2–14.0	8.04, 17.48
Bandwidth ΔE/E	$(2–5) \times 10^{-3}$	$\geq 10^{-2}$	2.5×10^{-4}
Photons/s (total)	2.5×10^{14}	2×10^{6}	$5 \times 10^{10}, 8 \times 10^{9}$
(on sample)			$5 \times 10^{6}, 8 \times 10^{5}$
Repetition rate (kHz)	0.12	2.0	1.0
Pulse duration (fs)	10–250	100	100
Timing jitter r.m.s. (fs)	200	<100	<100

(from ref. 1)

X-ray probes of ultrafast structural dynamics. Most studies of ultrafast structural dynamics make use of a pump-probe approach. A femtosecond optical excitation pulse induces a change of equilibrium structure and an ultrashort x-ray probe pulse of variable time delay maps the momentary structure of the excited sample. Resonant absorption of the pump pulse generates nonequilibrium populations of electronically and/or vibrationally excited states from which a structure changing process starts. Both excitations localized within a unit cell and delocalized electronic or propagating phonon excitations have been generated. For localized excitations, the fraction of excited unit cells in crystalline samples is typically less than 1 %. Interaction with nonresonant pump pulses induces a field-driven change, connected with a mixing of quantum states, i.e., a virtual excitation. In the probing step, interaction with the x-ray pulse should be in the linear regime of light-matter interaction and leave the momentary structure unchanged.
The time resolution of pump-probe experiments is determined by the duration of pump and probe pulses, the timing jitter between them, and the interaction geometry with the sample. Laser-based experiments in which both optical pump and x-ray probe are derived from a single laser system, as well as laser-driven slicing schemes at synchrotrons offer a timing jitter that is negligible compared to the respective pulse durations. A time resolution of ~100 fs has been demonstrated with both types of experiments. This issue is more critical when using FELs and an independent laser system for optical excitation. Here, the typical jitter has been of the order of 200 fs when averaging over many pump-probe events. Schemes for sorting individual pump-probe events in the recorded data set are being used for improving the time resolution [20]. This allows for shrinking the timing uncertainty to approximately 10 fs, close to the x-ray pulse duration.
Ultrafast structural dynamics have been probed via (i) nonresonant Bragg diffraction from (poly)crystalline materials, and (ii) resonant x-ray diffraction from crystalline materials with a correlated electron and/or spin system, including magnetic systems.

(i) Structural dynamics has been investigated by Bragg diffraction of ultrashort hard x-ray pulses from single crystals and/or crystalline powders [1-4]. The angular position of a particular Bragg reflection is determined by the condition $\mathbf{k'}$-$\mathbf{k}$=$\mathbf{G}_{hkl}$ where $\mathbf{k'}$ and $\mathbf{k}$ are the wavevectors of the scattered and the incoming x-ray pulse and $\mathbf{G}_{hkl}$ is the reciprocal lattice vector of the *(hkl)* set of lattice planes. The intensity of the Bragg peak $I_{hkl} \propto |F_{hkl}|^2$ is proportional to the square of the structure factor $F_{hkl} = |F_{hkl}|\exp(i\phi_{hkl})$ which represents the (spatial) Fourier transform of the electron density $\rho(\mathbf{r})$ in the unit cell of the crystal. While the time dependent position of a Bragg peak reflects the time-dependent spacing of lattice planes, its transient intensity gives insight into the redistribution of electronic charge. Time dependent structure is reconstructed from a sequence of diffraction patterns recorded for different pump-probe delays. Measuring intensity changes on a multitude of Bragg reflections allows for reconstructing time-dependent electron density maps.

(ii) Resonant x-ray diffraction [21] with a femtosecond time resolution has been implemented at LCLS [22–25]. In this method, the elastically scattered x-ray probe pulses are resonant to an atomic inner-shell transition of the system, in most cases in the soft x-ray regime. Under resonance conditions, the diffraction pattern contains - in addition to peaks from nonresonant Bragg diffraction - a resonantly enhanced contribution which gives rise to additional peaks. The resonant atomic scattering factor allows for separating different atomic species with their resonances at different transition energies. This 'chemical' sensitivity is particularly attractive when studying materials with large unit cells consisting of many atomic species. In general, resonant atomic scattering factors are tensorial quantities which reflect the ordering of electronic orbitals and the magnetic order of the system.

The sensitivity of femtosecond x-ray experiments, i.e., the smallest detectable change of diffracted intensity $\Delta I/I_0 = [I(t)-I_0]/I_0$ (I_0: diffracted intensity without excitation; t: delay time) depends on the stability of the pulse parameters such as intensity, temporal envelope, and overlap of pump and probe spots on the sample. A fundamental limitation is the photon counting statistics in the x-ray detection process. The counting shot noise leads to a relative uncertainty of $N^{-1/2}$ of the detected signal $\Delta I/I_0$ (N: number of x-ray photons counted by the detector). Counting shot noise is a major issue when using probe pulses with a comparably low x-ray flux from slicing or laser-driven plasma sources but there are sophisticated schemes to mitigate such noise [26]. The so far smallest signals of $\Delta I/I_0=10^{-3}$ have been measured with a 100 fs time resolution in powder diffraction experiments with a plasma source working at a 1 kHz repetition rate [27]. Typical integration times were several hours per delay position, realized by combining different data sets with a precise relative timing. Future FEL experiments with sufficient beam time at a high x-ray flux should reach a similar sensitivity with much shorter integration times.

Ultrafast x-ray diffraction

Real-space atomic motions and vibrational dynamics. Time-resolved x-ray diffraction has been applied to study different regimes of atomic motions. Appropriate optical excitation of a crystalline sample induces local atomic motions in the excited unit cells, typically less than 1% of all unit cells. Such motions have the character of optical phonon wavepackets and occur on a femtosecond time scale determined by the respective phonon frequencies. The related lattice elongations are a few percent of the lattice constant or chemical bond length at most. On a longer time scale, acoustic phonon and/or polariton propagation has been observed. This behavior is connected with mechanical strain propagation through the material, changing the lattice constants of a much larger fraction of unit cells. This results in changes of the shape and position of the rocking curve.

Both types of excitations eventually decay via dephasing and energy relaxation of the wavepacket. The population decay into other phonon modes via anharmonic coupling generates a quasi-equilibrium state at an elevated temperature and expands the lattice thermally.

There are different mechanisms for exciting lattice motions which have been reviewed in ref. [28]. The following excitation schemes have been applied in ultrafast x-ray diffraction:

(i) Displacive excitation: Electronic excitation via a bandgap results in a change of the electronic wavefunction and, thus, of the potential energy surface of vibrations/phonons which couple to the electronic transition. The minimum of the excited-state potential is shifted along the respective vibrational coordinate and the displaced initial vibrational wavefunction, a nonstationary wavepacket, moves along the vibrational coordinate, connected with coherent phonon elongations.

(ii) Raman excitation: Impulsive Raman excitation by a broadband femtosecond pulse generates a superposition of optical phonon or polariton eigenstates in the electronic ground state and a concomitant wavepacket motion. Excitation in the range of a dipole-allowed electronic transition enhances the Raman cross section resonantly as has been analyzed in detail in the time-domain wavepacket picture developed in Ref. [29].

(iii) Electronic excitation of metals and semiconductors allows for generating nonequilibrium conduction band electrons which thermalize into a hot (quasi-)Fermi distribution and transfer excess energy to the lattice via electron-phonon coupling. If the fully incoherent cooling process is fast compared to the phonon oscillation period, it can generate an 'impulsive' stress that drives coherent motions along low-frequency phonon coordinates.

Local oscillatory lattice motions in real-space are connected with a modulation of electronic charge density and, thus, of the x-ray structure factor. This gives rise to an intensity modulation of the corresponding Bragg peaks. The angular positions of the Bragg peaks remain unaffected as long as the size of the unit cells is unchanged. Phonon frequencies are directly mapped into the oscillation period of Bragg peak intensities whereas the extraction of spatial phonon amplitudes requires a quantitative analysis of the amplitude change of the structure factor. The subsequent

decay of the wavepacket and heating of the lattice induce an angular shift of Bragg peaks on time scales between a few and hundreds of picoseconds.
Ultrafast lattice motions and strain propagation have extensively been studied in both bulk and nanolayered crystalline materials. Examples can be found in [1-4,30-37].

Transient electron density maps. The structure factor $F_{hkl} = |F_{hkl}|\exp(i\phi_{hkl})$ of x-ray diffraction represents the Fourier transform of the electron density $\rho(\mathbf{r})$. This fundamental relation allows for deriving spatially resolved electron density maps from x-ray diffraction patterns containing a large number of Bragg peaks (*hkl*). Stationary x-ray Bragg and Laue diffraction have widely been applied to generate equilibrium charge density maps with up to picometer spatial resolution [38,39]. The recent implementation of ultrafast x-ray powder diffraction [40] which provides time resolved diffraction patterns consisting of up to 40 Debye Scherrer rings, has allowed for creating the first transient electron density maps with a time resolution of 100 fs. Such pioneering work has given new insight into field-driven electron relocations, the interplay of lattice and electron motions in ionic and ferroelectric crystals, as well as in elementary chemical processes such as hydrogen transfer [4,27,41-46].
The change of diffracted intensity $\Delta I_{hkl}/I_0 = [I_{hkl}(t)-I_0]/I_0$ integrated over a Debye Scherrer ring (*hkl*) is given by

$$\begin{aligned}\Delta I_{hkl}(t) &= M_{hkl} L P_{hkl}\left(|F_{hkl}(t)|^2 - \left|F^0_{hkl}\right|^2\right)\\ &= M_{hkl} L P_{hkl}\left(\left|\eta F^{ex}_{hkl}(t) + (1-\eta)F^0_{hkl}\right|^2 - \left|F^0_{hkl}\right|^2\right).\end{aligned}$$

Here, M_{hkl} is the multiplicity of the (*hkl*) diffraction ring and LP_{hkl} represents the Lorentz polarization factor. F^0_{hkl} and $F^{ex}_{hkl}(t)$ is the structure factor of the unit cell before and after excitation, respectively, and $\eta<<1$ is the fraction of excited unit cells. In calculating the quantity $|F_{hkl}(t)|^2$, the product term $F^{ex}_{hkl}(t)\times F^0_{hkl}$ describes the interference of x-rays diffracted from excited unit cells with those diffracted from unexcited unit cells. In other words, the strong x-ray component diffracted from unexcited unit cells serves as reference wave in heterodyning the much weaker x-rays diffracted from the small fraction η of modified unit cells.
Deriving electron density maps from the measured intensity changes $\Delta I_{hkl}(t)$ requires to know $F^{ex}_{hkl}(t)$ in amplitude and phase. For a powder of crystallites consisting of unit cells with inversion symmetry, the structure of excited crystallites averaged over all orientations in the powder is again inversion-symmetric. Thus, the initial phase $\phi_{hkl}=0$, is preserved. Neglecting terms quadratic in η, the following relation between the change in the structure factor $\Delta F_{hkl}(t) = F^{ex}_{hkl}(t)-F^0_{hkl}$ and $\Delta I_{hkl}(t)$ is derived:

$$\eta\Delta F_{hkl}(t) = \eta\left(F^{ex}_{hkl}(t) - F^0_{hkl}\right) = \frac{\Delta I_{hkl}(t)}{2M_{hkl} L P_{hkl} F^0_{hkl}}$$

Here, all quantities on the r.h.s are known, allowing for determining $\Delta F_{hkl}(t)$. The latter provide the change $\Delta\rho(\mathbf{r}, t)$ of electronic charge density via the Fourier series

$$\eta\Delta\rho(\mathbf{r},t) = \frac{\eta}{|\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}|}\sum_{h,k,l}\Delta F_{hkl}(t) \times \cos[2\pi(h\,\mathbf{a}^* + k\,\mathbf{b}^* + l\,\mathbf{c}^*)\mathbf{r}]$$

a*, **b***, and **c*** are the reciprocal lattice vectors of the lattice vectors **a**, **b**, and **c** of the unit cell. The spatial resolution and, thus, the accuracy of the generated electron density maps increases with the number of Debye Scherrer rings (*hkl*) recorded up to a maximum diffraction angle θ_{max}. Deriving transient charge densities for crystalline systems without inversion symmetry requires more sophisticated methods [27], including iterative numerical procedures such as, e.g., the Maximum Entropy Method (MEM) [4,47-49].

In the lecture, different examples of transient charge density maps will be discussed. In polar or ionic systems, in particular ferroelectrics, they reveal a subtle interplay of small-amplitude lattice motions and large-amplitude charge relocations. Such behavior is characteristic for so-called soft-modes [50] which play a key role in para- to ferroelectric phase transitions. With the time-dependent charge density maps at hand, one can go one step further and derive the time-dependent macroscopic electric polarization of the excited material from the microscopic charge distributions, in this way solving a long-standing problem of condensed matter physics [46].

References

[1] T. Elsaesser, M. Woerner, J. Chem. Phys. 140, 020901 (2014).
[2] T. Elsaesser and M. Woerner, Acta Cryst. Sec. A 66, 168 (2010).
[3] C. v. Korff Schmising, M. Bargheer, M. Woerner, T. Elsaesser, Z. Krist. 223, 283 (2008).
[4] M. Woerner *et al.*, Faraday Disc. 171, 373 (2014).
[5] R.W. Schoenlein *et al.*, Science 274, 236 (1996).
[6] S. Khan et al., Phys. Rev. Lett. 97, 074801 (2006).
[7] P. Beaud et al., Phys. Rev. Lett. 99, 174801 (2007).
[8] S. Khan, J. Mod. Opt. 55, 3469 (2008).
[9] P. Emma *et al.*, Nature Photon. 4, 641 (2010).
[10] T. Ishikawa *et al.*, Nature Photon. 6, 540 (2012).
[11] I. A. Vartanyants *et al.*, Phys. Rev. Lett. 107, 144801 (2011).
[12] J. Amann *et al.*, Nature Photon. 6, 693 (2012).
[13] M. M. Murnane, H. C. Kapteyn, M. D. Rosen, R. W. Falcone, Science 251, 531 (1991).
[14] F. Brunel, Phys. Rev. Lett. 59, (1987) 52.
[15] N. Zhavoronkov *et al.*, Opt. Lett.30, 1737 (2005).
[16] F. Zamponi *et al.*, Appl. Phys. A 96, 51(2009).
[17] J. Weisshaupt *et al.*, Nature Photon. 8, 927 (2014).
[18] J. Weisshaupt *et al.*, Struct. Dyn. 2, 024102 (2015).
[19] L. von Grafenstein, Optics Letters 42, 3796 (2017).
[20] M. Harmand *et al.*, Nature Photon. 7, 215 (2013).
[21] J.-L. Hodeau *et al.*, Chem. Rev. 101, 1843 (2001).
[22] S. L. Johnson *et al.*, Phys. Rev. Lett. 108, 037203 (2012).
[23] C. E. Graves *et al.*, Nature Mater. 12, 293 (2013).

[24] Y. D. Chuang *et al.*, Phys. Rev. Lett. 110, 127404 (2013).
[25] S. de Jong *et al.*, Nature Mater. 12, 882 (2013).
[26] M. Holtz *et al.*, Struct. Dyn. 4, 054304 (2017).
[27] F. Zamponi, P. Rothhardt, J. Stingl, M. Woerner, T. Elsaesser, Proc. Nat. Acad. Sci. USA 109, 5207 (2012).
[28] M. Först and T. Dekorsy, in : Coherent vibrational dynamics, S. DeSilvestri, G. Cerullo, G. Lanzano (Eds.), CRC Press (Taylor and Francis Group), London, p. 130 (2008).
[29] S.Y. Lee, E. J. Heller, J. Chem. Phys. 71, 4777 (1979).
[30] K. Sokolowski-Tinten *et al.*, Nature 422, 287 (2003).
[31] D. M. Fritz et al., Science 315, 633 (2007).
[32] S. L. Johnson, E. Vorobeva, P. Beaud, C.J. Milne, and G. Ingold, Phys. Rev. Lett. 103, 205501 (2009).
[33] Y. Giret, A. Gell, B. Arnaud, Phys. Rev. Lett. 106, 155503 (2011).
[34] M. Bargheer *et al.*, Science 306, 1771 (2004).
[35] C. v. Korff Schmising *et al.*, Phys. Rev. Lett 98, 257601 (2007).
[36] C. v. Korff Schmising *et al.,* Phys. Rev. B 78, 060404(R) (2008).
[37] D. Schick *et al.*, Phys. Rev. Lett. 110, 095502 (2013).
[38] P. Coppens, X-Ray Charge Densities and Chemical Bonding, Oxford University Press 1997.
[39] C. Gatti, P. Macchi (Eds.), Modern Charge Density Analysis, Springer, Heidelberg 2012.
[40] F. Zamponi, Z. Ansari, M. Woerner, T. Elsaesser, Opt. Express 18, 947 (2010).
[41] M. Woerner *et al.,* J. Chem. Phys. 133, 064509 (2010).
[42] F. Zamponi *et al.*, Phys. Chem. Chem. Phys. 14, 6156 (2012).
[43] J. Stingl *et al.,* Phys. Rev. Lett. 109, 147402 (2012).
[44] V. Juvé *et al.*, Phys. Rev. Lett. 111, 217401 (2013).
[45] B. Freyer *et al.*, J. Chem. Phys. 138, 144504 (2013).
[46] C. Hauf et al., Struct. Dyn. 5, 024501 (2018).
[47] E. Jaynes, Phys. Rev. 108, 171 (1957).
[48] C. Gilmore, Acta Cryst. A 52, 561 (1996).
[49] S. Smaalen, L. Palatinus, and M. Schneider, Acta Cryst. A 59, 459 (2003).
[50] W. Cochran, Adv. Phys. 9, 387 (1960).

Quantum Crystallography and Crystal Engineering: Experimental lattice energies from X-ray diffraction data?

Mark A. Spackman
School of Molecular Sciences, University of Western Australia, Perth, WA, Australia
mark.spackman@uwa.edu.au

Introduction

At this stage of the School you should have an excellent idea of what *quantum crystallography* is, and the various approaches – experimental and computational – that can be used to obtain information about the structure, bonding, dynamics and properties of crystalline solids. In this lecture I would like to focus attention on some of the ways in which quantum crystallography, particularly in the form of charge density analysis of X-ray diffraction data, is being applied to problems in *crystal engineering*.

Although not limited to molecular crystals, the term *crystal engineering* is generally understood to mean something like "the understanding of intermolecular interactions in the context of crystal packing and the utilization of such understanding in the design of new solids with desired physical and chemical properties",[1] and this is how the term will be understood in this lecture. Crystal engineering is often invoked in the introduction to charge density studies on molecular crystals, and the following excerpts highlight the way in which these applications of charge density analysis are typically focused on polymorphism, and related topics associated with compounds of pharmaceutical relevance:

"In recent years, the occurrence of polymorphism in molecular crystals has received considerable attention, especially from the drug design and crystal engineering viewpoint".[2]

"The study of polymorphism in piroxicam provides us knowledge on the behaviour of zwitterionic molecules and therefore gives us potential insights into new methods available which may benefit efforts in crystal engineering".[3]

"The utility of cocrystals in crystal engineering specifically for improving active pharmaceutical ingredients (APIs) has recently refocused the scientific spotlight onto the field of crystallography".[4]

"Crystal engineering refers to the rational design of solids by inducing the reproducible formation of weak interactions between pairs of functional groups. This field has recently undergone a resurgence because of increased interest from pharmaceutical companies and research institutions. This coincides with a recent reduction in the number of novel drugs approved by the FDA and other global regulatory agencies relative to the early 1990s".[5]

The excellent 2014 review article "Charge density analysis for crystal engineering", co-authored by one of the organisers of this School, has covered almost all of the content that might be expected to appear in a lecture with the title at the top of this page. But I would like to do something quite different here, and highlight a topic that was not covered in much detail in that review: I would like to shine a forensic light on recent attempts to derive quantitative estimates of intermolecular interaction energies, and especially lattice energies, from the multipolar modelling of X-ray diffraction data.

The intermolecular interaction energy

Before discussing actual examples, it is important to establish precisely what we mean by this interaction energy, its relationship with the charge distributions of the interacting molecules, especially in the context of experimental molecular electron densities. For the past 50 years or so (see ref. [6] and references therein) the interaction energy between two molecules has been conveniently expressed as a sum of several discrete terms, the main ones being electrostatic, polarisation (or induction), dispersion and exchange-repulsion:

$$E_{tot} = E_{ele} + E_{pol} + E_{dis} + E_{rep} \quad 1)$$

The first term is simply the classical Coulombic interaction between the two unperturbed molecular charge distributions (electrons and nuclei), and it can be positive (= de-stabilising) or negative (= stabilising). The second is also electrostatic in nature, being the energy lowering associated with the perturbation (polarisation) of the electron density of each molecule by the other, and is always negative. The third term arises from non-classical effects and is also always negative, while the last term is always positive (as its name implies), and arises from the overlap of the two molecular wavefunctions.

There are several important points to note when using this expression in experimental quantum crystallography:

1. Only the first two terms are accessible from the electron density of a molecule extracted from its crystal environment; this is kind of obvious.
2. A molecular electron distribution extracted by modelling of experimental data necessarily reflects the perturbation due to its surrounding crystal environment.
3. There are many possible model electron densities that will provide only slightly different fits to the experimental observations, so any one that is chosen as the 'best' necessarily includes the effects of systematic and random errors.
4. Because of 2. it is commonly assumed (and often stated emphatically) that the calculation of the electrostatic energy between two perturbed molecules yields the first two terms above, E_{ele}(polarized) = E_{ele} + E_{pol}. This is not true. Careful analysis[7] shows instead that it overestimates the 'true' electrostatic energy, and in fact E_{ele}(polarized) is closer to E_{ele} + 2E_{pol}.
5. E_{dis} and E_{rep} are typically estimated via an atom-atom potential, usually of exp-6 form, and commonly those due to Spackman[8] or Williams and Cox.[9] This is an approximation, as the atom-atom representation of dispersion and repulsion energies depends intimately on the way in which the electrostatic terms are computed. In short, those atom-atom potentials are not compatible with the electrostatic energy obtained from experimental charge distributions.
6. Because of 3. estimating the uncertainty associated with the experimentally-derived electrostatic energy is essential. Without any idea of the likely magnitude of this

uncertainty, any comparison with other results, experimental or theoretical, is far less meaningful.

Because the electrostatic energy is clearly the most important energetic quantity that can be derived from experimental quantum crystallography, we also need to recognise the different ways in which this term has been calculated – now, as well as in the past. To see these differences most clearly, note that the Coulombic interaction energy between two molecules A and B, with charge densities $_{A}(\mathbf{r}_A)$ and $_{B}(\mathbf{r}_B)$, is

$$E_{ele} = \iint \rho_A(\boldsymbol{r}_A)\rho_B(\boldsymbol{r}_B) \mid \boldsymbol{r}_A - \boldsymbol{r}_B \mid^{-1} \mathrm{d}\boldsymbol{r}_A \mathrm{d}\boldsymbol{r}_B \tag{2}$$

These charge densities include point nuclear charges and electron densities, and they can be partitioned into atom-centred pseudoatoms, each consisting of a spherical atomic term (including a nucleus) and a deformation term:

$$\begin{aligned} \rho_A &= \sum_{i \in A} \{\rho_{A,i}^{atomic} + \Delta\rho_{A,i}\} \\ &= \rho_A^{promolecule} + \Delta\rho_A \end{aligned} \tag{3}$$

From this the product of molecular charge densities in eq. (1) can be expanded to give

$$\rho_A\rho_B = \rho_A^{pro}\rho_B^{pro} + (\rho_A^{pro}\Delta\rho_B + \rho_B^{pro}\Delta\rho_A) + \Delta\rho_A\Delta\rho_B$$

and substituting this expansion into eq. (2) results in an expression like this:

$$E_{ele} = E_{ele}^{pro-pro} + E_{ele}^{pro-def} + E_{ele}^{def-def} \tag{4}$$

Early attempts to calculate electrostatic energies from experimental electron densities approximated the electron distribution by a sum of atom-centred multipole moments (point charges, dipoles, quadrupoles etc. on each atom), and used expressions due to Buckingham[10] to compute the electrostatic energy. But the spherical atomic charge densities have zero multipole moments, so this approximation amounts to ignoring the first two terms in eq. (4). Numerous examples are available where this approach was used,[7b,11] but the promolecule contribution to the electrostatic energy in eq. (4), $E_{ele}^{pro-pro}$, is always negative at intermolecular separation distances,[12] and typically in the range –10 to –70 kJ mol^{-1} for nearest neighbour intermolecular interactions.[13] As such it is essential to include that contribution, and several studies have been reported that corrected the multipole-moment term in eq. (4) by adding a close approximation to $E_{ele}^{pro-pro} + E_{ele}^{pro-def}$.[14]

The accurate calculation of electrostatic interaction energies from experimental charge densities, eq. (2), was significantly advanced by the publication of the combined exact potential and multipole method (EP/MM),[15] which "combines numerical quadrature evaluation of integrals involving the electron density and potentials for short-range pseudoatom-pseudoatom interactions with the standard Buckingham-type multipole approximation for long-range interactions". Incorporation of this method into *XD*[16] has resulted in the reporting of "experimental" intermolecular interaction energies – and lattice energies – as outcomes of a

charge density analysis of X-ray diffraction data, and these results are the subject of this lecture (and the reason for the question in the title). Numerical calculation of accurate electrostatic energies, eq. (2), can also be performed with *VMoPro*, a properties visualization part of the *MoPro* software,[17] but I am aware of only one report[18] of experimental electrostatic energies using that software, and none of lattice energies.

The lattice energy

Rather than discuss numerous examples of interaction energies, for molecular pairs with different orientations, arising from something like twenty recent charge density studies, I am going to focus on the sum of those energies – the lattice energy. This is not simply a matter of convenience (although one quantity per crystal structure makes for a more succinct discussion). The logic here is that the lattice energy is closely related to a thermodynamic quantity – the sublimation enthalpy – for which experimental measurements are available for a very large number of crystalline materials.[19] And of course lattice energies can aid in understanding the relative stability of polymorphs and co-crystals.

For the purpose of this discussion let's assume that we can calculate E_{tot} in eq. (1) from a pseudoatom model (or even X-ray constrained wavefunctions) for any pair of molecules A and B in a crystal. This requires accurate calculation of E_{ele}(polarized), as well as some choice of atom-atom potential to approximate the dispersion and repulsion energies. Then the lattice energy can be computed via a sum of pairwise energies:

$$E_{lat} = \frac{1}{2} \sum_{R_{AB}<R} E_{tot}^{AB} . \qquad (5)$$

This can be related to the sublimation enthalpy by taking into account the electronic relaxation energy associated with the crystal to gas transition, as well the difference between vibrational energies for the two phases,

$$\Delta H_{sub}(T) = -E_{lat} + \Delta E_{rel} + \Delta E_{vib} + 4RT , \qquad (6)$$

and this can be (reasonably well[20]) approximated by

$$\Delta H_{sub}(T) = -E_{lat} + \Delta E_{rel} - 2RT . \qquad (7)$$

The important point here is that for crystals where little molecular geometry change occurs between crystal and gas phase, the sublimation enthalpy is a positive quantity, whose magnitude at room temperature is ~5 kJ mol^{-1} greater than that of the lattice energy. This means we have a very useful primary benchmark against which to validate lattice energies (and by inference, pairwise interaction energies) derived from X-ray diffraction data – bearing in mind of course the inherent experimental errors of ~5 kJ mol^{-1} or more in these thermodynamic measurements.

CrystalExplorer model energies

Although the latest compilations of phase transition enthalpies are comprehensive, sublimation enthalpies are not always available for the molecular crystals of interest. However, we can use CE-B3LYP lattice energies as secondary benchmarks where sublimation data are lacking. And we can also potentially use the individual energy components in the CE-B3LYP model to tell us something about the "experimental" electrostatic energy.

CE-B3LYP model energies[21] incorporated in *CrystalExplorer*[22] utilise B3LYP/6-31G(d,p) quantum mechanical charge distributions for unperturbed monomers, and separate the interaction energy between pairs of molecules or ions into electrostatic, polarization, dispersion and exchange-repulsion terms, much like eq. (1):

$$E_{\text{tot}} = k_{\text{ele}}E_{\text{ele}} + k_{\text{pol}}E_{\text{pol}} + k_{\text{dis}}E_{\text{dis}} + k_{\text{rep}}E_{\text{rep}}. \quad (8)$$

Here E_{ele} is the same as in eq. (1), and E_{pol} is a sum of terms of the kind $-\frac{1}{2}\alpha F^2$, where are isotropic polarizabilities and *F* the electric field magnitude computed at each atomic nucleus resulting from the charge distribution of the other monomer. Details of the other terms are given elsewhere[21,23] and optimum values of the scale factors k_{ele}, etc. in eq. (8) were determined by calibration against counterpoise-corrected B3LYP-D2/6-31G(d,p) interaction energies for a large number of molecule/ion pairs. The mean absolute deviation (MAD) of these CE-B3LYP model energies from the DFT benchmark values is 2.4 kJ mol^{-1} for energies of molecule/ion pairs that span a range of 3.75 MJ mol^{-1}.[23] The same model energies yield lattice energies with a MAD from benchmark values of only 6.6 kJ mol^{-1}.[24]

Comparison between experimental and CE-B3LYP lattice energies

Table 1 compares lattice energies reported using the EP/MM approach in *XD* for the electrostatic energy (and atom-atom potentials, as indicated) with CE-B3LYP lattice energies based on the crystal structures tabulated. Experimental sublimation enthalpies are also given for a subset of results.

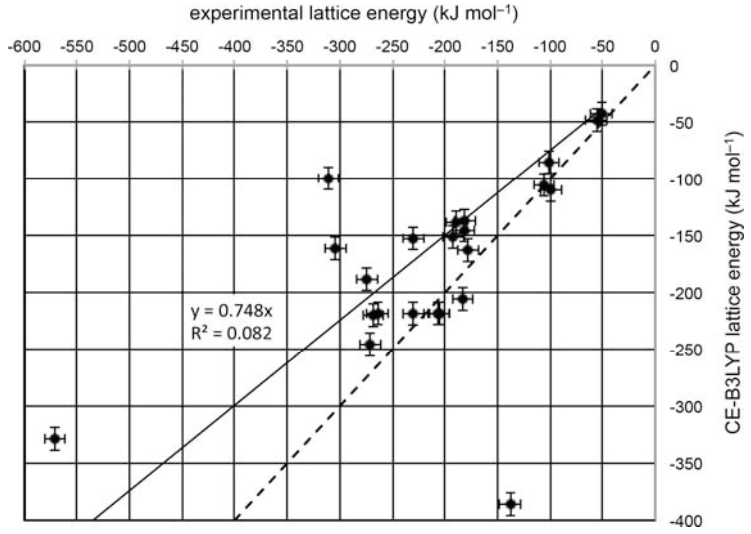


Figure 1. Experimental lattice energies compared with CE-B3LYP results from Table 1. The solid line is a linear regression, and the dashed line has unit slope, passing through the origin. Nominal uncertainties of ±10 kJ mol^{-1} are indicated for both quantities.

From Table 1 we see that CE-B3LYP lattice energies are consistent with sublimation enthalpies, bearing in mind the inherent errors in both quantities, and the fact that the relaxation energy has been ignored in these CE-B3LYP estimates.

Figure 1 shows that (superficially at least) the experimental and CE-B3LYP lattice energies are poorly correlated. Almost all experimental lattice energies are greater in magnitude than CE-B3LYP results (i.e. they lie above the dashed line in the figure) – and sometimes much greater. The exceptions are forms I, II and IV of sulfathiazole, and pyrazinamide (and for which the difference is likely to be within experimental error), and form II of the $(4HBA)_2(44BP)$ co-crystal. The difference between the results for the $(4HBA)_2(44BP)$ co-crystal form II is 248 kJ mol^{-1}, well outside any reasonable experimental error, and it merits closer inspection as the difference for 4,4'-bipyridine (44BP) in the same publication[4] is 212 kJ mol^{-1}, but of opposite sign. A plausible explanation is that these values do not properly reflect the stoichiometry of these two crystals. 44BP (HIQWEJ03) has two molecules in the asymmetric unit, and the chemical formula for 44BP is reported incorrectly in the original work as $C_{20}H_{16}N_4$, rather than $C_{10}H_8N_2$. But it makes no sense to report a lattice energy per two identical molecules, as this ignores the connection with the sublimation process, where the molecules are indistinguishable in the gas phase.

Table 1. Lattice energies (kJ mol^{-1}) estimated from X-ray diffraction data compared with CE-B3LYP model energies and experimental sublimation enthalpies, where available.[†]

compound	exp-6 potential	CSD refcode	experiment E_{lat}	CE-B3LYP E_{lat}	$_{sub}H$
dimethyl 3-(4-nitrophenyl) aziridine-2,2-dicarboxylate [25]	W&C	WIGJAY01	–178	–163	
dimethyl 2-(4-nitrophenyl) ethene-1,1-dicarboxylate [25]	W&C	XOCYOE	–182	–145	
(±)-8'-benzhydrylideneamino-1,1'-binaphthyl-2-ol [26]	W&C	WUFJOW01	–274	–188	-
oxirane [27]	W&C	DUFBOV11	–51	–43	41 [28]
dimethyl oxirane-2,3-dicarboxylate [27]	W&C	EYIJAY	–106	–105	-
dimethyl 3-(4-nitrophenyl) oxirane-2,2-dicarboxylate [27]	W&C	XOCYUK	–192	–151	-
oxirane-2,2,3,3-tetracarbonitrile [27]	W&C	TCYNEO04	–101	–85	-
sulfathiazole form I [29]	W&C	SUTHAZ43	–183	–206	-
sulfathiazole form II [29]	W&C	SUTHAZ30	–205	–218	-
sulfathiazole form III [29]	W&C	SUTHAZ33	–230	–219	-
sulfathiazole form IV [29]	W&C	SUTHAZ36	–206	–218	-
sulfathiazole form V	-	SUTHAZ05	-	–202	-
pyrazinamide [30]	S	PYRZIN22	–99	–110	116(4) [19a]
rubrene (100 K) [31]	W&C	QQQCIG17	–264	–218	181 [32]
rubrene (20 K) [31]	W&C	QQQCIG23	–268	–220	181 [32]
carbamazepine form III [33]	?	CBMZPN22	–230	–152	-
-piroxicam [3]	?	BIYSEH14	–304	–161	-
piroxicam monohydrate [3]	?	CIDYAP05	–571	–329	-
(4HBA) [4]	S	JOZZIH01	–181	–137	117(4) [19a]
4,4'-bipyridine (44BP) [4]	S	HIQWEJ03	–311	–99	106(3) [19a]
$(4HBA)_2$(44BP) co-crystal form II [4]	S	EPUPUB03	–138	–386	-
$(4HBA)_2$(44BP) co-crystal form I	-	EPUPUB	-	–396	-
paracetamol (PCM) [5]	?	HXACAN07	–189	–138	126(6) [19a]
(PCM)(44BP) co-crystal [5]	?	MUPQAP	–271	–246	-
benzene [25]	W&C	[‡]	–56	–48	45[32]

[†] All experimental results employed the EP/MM approach in *XD* for the electrostatic energy, and atom-atom potentials, usually from either Williams and Cox[9] (W&C) or Spackman[8] (S), for dispersion and exchange-repulsion energies. The experimental values tabulated are unchanged from those reported in the original publications. CE-B3LYP lattice energies are based on the crystal structures indicated.

[‡] Based on X-ray data and crystal structure reported in ref. [34].

The co-crystal form II (EPUPUB03) has formula $C_{10}H_8N_2.2(C_7H_6O_3)$, but the asymmetric unit consists of one molecule of 4HBA and one-half of a 44BP molecule.

With reference to the sublimation process, one mole of the co-crystal will generate two moles of molecules of 4HBA and one of 44BP. This is rather confusing, but it makes more sense if we consider the hypothetical chemical process:

$$44BP(s) + 2 \times 4HBA(s) \rightarrow (4HBA)_2(44BP)(s) \qquad \Delta E_{lat}$$

Assuming no changes in molecular geometries, we would expect this process to be accompanied by a lowering in intermolecular energy, i.e. $E_{lat} < 0$.

E_{lat}(form II) = $E_{lat}((4HBA)_2(44BP))$ – $[2E_{lat}(4HBA) + E_{lat}(44BP)]$

= –386 –[2(–137) + –99)] = –13 kJ mol^{-1} using CE-B3LYP results

= –138 –[2(–181) + –311)] = +535 kJ mol^{-1} using experimental results

This makes it clear that the experimental lattice energies must be incorrect.

We can push this discussion a little further by using CE-B3LYP energies to compare the relative stabilities of forms I and II of the $(4HBA)_2(44BP)$ co-crystal.

ΔE_{lat}(form I) = –396 –[2(–137) –99)] = –23 kJ mol^{-1} using CE-B3LYP results from Table 1, suggesting that form I is more stable than form II, and this agrees nicely with experimental observations.[35] A similar analysis of the paracetamol (PCM) co-crystal with 44BP yields ΔE_{lat} = +229 kJ mol^{-1} from experiment,[5] but –9 kJ mol^{-1} using CE-B3LYP results.

Comments on the experimental electrostatic energy

The systematic overestimate of the lattice energy from experimental electron distributions provides a clue to a possible origin in the electrostatic term. As noted earlier, E_{ele}(polarized) is actually closer to the sum of $E_{ele} + 2E_{pol}$ than it is to $E_{ele} + E_{pol}$, and certainly greater in magnitude than E_{ele}. Fortunately, most publications of experimental lattice energies also report the electrostatic term, and we can use these results to explore this overestimate – see Figure 2.

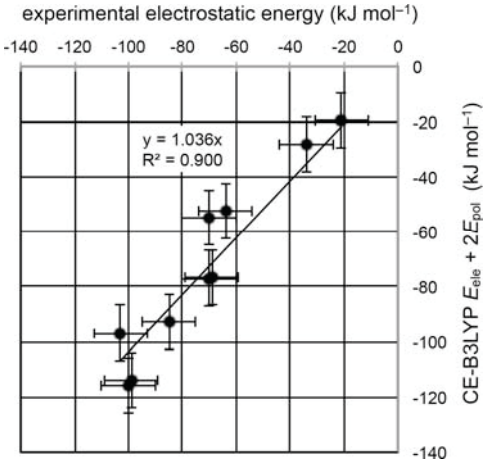


Figure 2. Electrostatic components of ten experimental lattice energies[25,27,30-31] compared with the CE-B3LYP sum of scaled components $E_{ele} + 2E_{pol}$. Nominal uncertainities of ±10 kJ mol^{-1} are indicated for both quantities.

In this figure I have chosen to use the most reliable (in my opinion) experimental estimates, and exclude results where the stoichiometry made little sense (as discussed above). Error estimates in the experimental quantities were not reported, but they are likely to be ~10 kJ mol^{-1}. The line of best fit passing through the origin suggests that the experimental electrostatic energy correlates strongly with the sum of $E_{ele} + 2E_{pol}$ from CE-B3LYP lattice energies. Note that the figure plots the sum of scaled CE-B3LYP energy terms, giving a MAD of 9 kJ mol^{-1}, but unscaled terms yield a similar result of 11 kJ mol^{-1}. This suggests that the CE-B3LYP E_{pol} energy is a rough measure of the overestimate in the true electrostatic energy when calculated using polarized charge distributions. Representative values of E_{pol} are ~2 kJ mol^{-1} for benzene, ~9 kJ mol^{-1} for rubrene, ~11 kJ mol^{-1} for oxirane-2,2,3,3-tetracarbonitrile and ~22 kJ mol^{-1} for dimethyl 3-(4-nitrophenyl)aziridine-2,2-dicarboxylate.

Comments on the atom-atom potentials

This analysis of electrostatic energies can partly explain the apparent experimental overestimate of lattice energies (and a possible correction for that), but we also need to examine the different atom-atom potentials used to provide corresponding estimates of dispersion and exchange-repulsion energies.

The "Williams & Cox" atom-atom (nonbonded) potentials were originally derived by fitting to oxohydrocarbon[36] and azahydrocarbon[9] crystal structures, using atomic point charges to represent the electrostatic energy. Different refinements of these exp-6 potentials have been used extensively in crystal structure prediction by Sally Price, Graeme Day and co-workers, and a recent publication[37] summarizes the differences between the versions, and makes it clear that different versions are appropriate for different descriptions of the electrostatic energy: distributed atomic multipoles *vs* atomic point charges fitted to the molecular electrostatic potential. The precise "Williams & Cox" atom-atom potential terms used in *XD* are those for H, C, N and O reported in refs. [9,36].

In the atom-atom potentials that I reported in 1986,[8] the exchange-repulsion terms were obtained using a rather simple density functional approach based on spherical atomic (promolecular) charge distributions. Empirical C_6 dispersion coefficients available at that time were added to yield a consistent set of exp-6 potentials for pairs atoms up to and including Br**.** These were successfully applied to obtain geometries, interaction energies, force constants and vibrational frequencies of hydrogen bonded dimers of small linear molecules,[38] but it is important to recognise some of the assumptions in that work that may have been overlooked:

1. The expression for the exchange-repulsion energy in ref. [8] actually includes the electrostatic interaction between spherical atomic charge densities, $E_{ele}^{pro-pro}$.
2. Because of 1. only the use of point atomic multipoles to compute the electrostatic component of the total interaction energy makes sense, along with an estimate of $E_{ele}^{pro\text{-}def}$ (labelled E_{pen} in ref. [38]).

3. The repulsion part of the exp-6 potential for a hydrogen atom involved in a hydrogen bond is set to zero.

Points 1. and 2. above imply that the "Spackman" exp-6 potentials should not be used in combination with experimental electrostatic energies computed with the EP/MM approach (or any other essentially exact method). And point 3. is problematic for general application to molecular crystals, as it requires making assumptions about which atoms are hydrogen bonded, and which are not.

From this analysis is seems that neither the "Williams & Cox" or "Spackman" exp-6 potentials are ideally suited to providing dispersion and exchange-repulsion energies that complement experimental electrostatic energies calculated by accurate methods. This then clearly identifies a need to develop an appropriate set of nonbonded potentials for this purpose. I am aware of one attempt to modify the "Spackman"

exp-6 potentials by fitting to *ab initio* repulsion energies and experimental dipole oscillator strength distributions (for the C_6 dispersion coefficients), but the results have not been reported in detail.[14c]

Concluding remarks

These lecture notes are necessarily brief, and relatively succinct, but they will be discussed in greater depth in the lecture. The take home messages are relatively straightforward:

- The calculation of intermolecular interaction energies – and lattice energies – as an outcome of experimental charge density analysis *is still in its infancy*.
- Future calculations of this kind must acknowledge that the electrostatic energy calculated between polarized molecules does not simply "include the effects of polarization"; it actually overestimates them, sometimes by a large margin.
- Improvements must be made to the nonbonded atom-atom potentials used to approximate the dispersion and repulsion energies. A first step in this direction may be to use dispersion models such as the simpler D2 model due to Stefan Grimme.[39]
- Realistic estimates of the experimental error in the computed experimental electrostatic energy are not just desirable – they are essential in order to obtain results that are more than just numbers in a table.
- Ignoring basic thermochemical knowledge in presenting and discussing the experimental energies is unacceptable. The relationship of lattice energies to experimental sublimation enthalpies, and the widespread availability of the latter quantities, needs to be exploited as much as possible. And if experimental values make little sense compared with PIXEL or CE-B3LYP results, that discrepancy is surely telling us something important.

These notes have focused on multipole-refined model electron distributions, which have well known limitations that give rise to unavoidable systematic errors (and which can be hard to quantify). Would the situation described above be improved if instead we used X-ray constrained wavefunctions to compute all quantum mechanical-based energy terms

(electrostatic, polarization, exchange-repulsion), and couple those with a modern model of the dispersion energy?

References

[1] G. R. Desiraju, (Ed.) Crystal Engineering: The Design of Organic Solids, Elsevier, Amsterdam, 1989.

[2] P. Munshi, C. Jelsch, V. R. Hathwar, T. N. Guru Row, *Cryst. Growth Des.* **2010**, *10*, 1516.

[3] F. Lai, J. J. Du, P. A. Williams, L. Varadi, D. Baker, P. W. Groundwater, J. Overgaard, J. A. Platts, D. E. Hibbs, *Phys. Chem. Chem. Phys.* **2016**, *18*, 28802.

[4] J. J. Du, S. A. Stanton, P. A. Williams, J. A. Ong, P. W. Groundwater, J. Overgaard, J. A. Platts, D. E. Hibbs, *Cryst. Growth Des.* **2018**, *18*, 1786.

[5] J. J. Du, F. Lai, L. Váradi, P. A. Williams, P. W. Groundwater, J. A. Platts, D. E. Hibbs, J. Overgaard, *Crystals* **2018**, *8*, 46.

[6] K. Kitaura, K. Morokuma, *Int. J. Quantum Chem.* **1976**, *10*, 325.

[7] a) A. J. Stone, *The Theory of Intermolecular Forces*, Clarendon Press, Oxford, **1996**; b) K. Y. Suponitsky, V. G. Tsirelson, D. Feil, *Acta Crystallogr. A* **1999**, *55*, 821.

[8] M. A. Spackman, *J. Chem. Phys.* **1986**, *85*, 6579.

[9] D. E. Williams, S. R. Cox, *Acta Crystallogr. B* **1984**, *40*, 404.

[10] A. D. Buckingham, *Adv. Chem. Phys.* **1967**, *12*, 107.

[11] a) Z. Berkovitch-Yellin, L. Leiserowitz, *J. Am. Chem. Soc.* **1980**, *102*, 7677; b) Z. Berkovitch-Yellin, L. Leiserowitz, *J. Am. Chem. Soc.* **1982**, *104*, 4052; c) M. A. Spackman, H.-P. Weber, B. M. Craven, *J. Am. Chem. Soc.* **1988**, *110*, 775; d) Y. A. Abramov, A. Volkov, G. Wu, P. Coppens, *Acta Crystallogr. A* **2000**, *56*, 585; e) Y. A. Abramov, A. Volkov, G. Wu, P. Coppens, *J. Phys. Chem. B* **2000**, *104*, 2183; f) Y. A. Abramov, A. Volkov, P. Coppens, *J. Mol. Struct.* **2000**, *529*, 27; g) R. Destro, P. Roversi, M. Barzaghi, R. E. Marsh, *J. Phys. Chem. A* **2000**, *104*, 1047; h) X. Li, G. Wu, Y. A. Abramov, A. V. Volkov, P. Coppens, *Proc. Nat. Acad. Sci. USA* **2002**, *99*, 12132; i) D. E. Hibbs, C. J. Austin-Woods, J. A. Platts, J. Overgaard, P. Turner, *Chem. Eur. J.* **2003**, *9*, 1075; j) D. E. Hibbs, J. R. Hanrahan, M. B. Hursthouse, D. W. Knight, J. Overgaard, P. Turner, R. O. Piltz, M. P. Waller, *Org. Biomol. Chem.* **2003**, *1*, 1034; k) D. E. Hibbs, J. Overgaard, C. Gatti, T. W. Hambley, *New J. Chem.* **2003**, *27*, 1392; l) J. Overgaard, D. E. Hibbs, *Acta Crystallogr. A* **2004**, *60*, 480; m) P. Munshi, T. N. Guru Row, *Cryst. Growth Des.* **2006**, *6*, 708.

[12] M. A. Spackman, E. N. Maslen, *J. Phys. Chem.* **1986**, *90*, 2020.

[13] M. A. Spackman, *Chem. Phys. Lett.* **2006**, *418*, 158.

[14] a) R. Soave, M. Barzaghi, R. Destro, *Chem. Eur. J.* **2007**, *13*, 6942; b) R. Destro, R. Soave, M. Barzaghi, *J. Phys. Chem. B* **2008**, *112*, 5163; c) R. Destro, E. Sartirana, L. Loconte, R. Soave, P. Colombo, C. Destro, L. Lo Presti, *Cryst. Growth Des.* **2013**, *13*, 4571; d) G. Saleh, R. Soave, L. Lo Presti, R. Destro, *Chem. Eur. J.* **2013**, *19*, 3490.

[15] A. Volkov, T. Koritsansky, P. Coppens, *Chem. Phys. Lett.* **2004**, *391*, 170.

[16] a) T. Koritsánszky, P. Mallinson, P. Macchi, A. Volkov, C. Gatti, T. Richter, L. Farrugia, XD2006. A Computer Program Package for Multipole Refinement, Topological Analysis of Charge Densities and Evaluation of Intermolecular Energies from Experimental or Theoretical Structure Factors, 2007. ; b) A. Volkov, P. Macchi, L. J. Farrugia, C. Gatti, P. Mallinson, T. Richter, T. Koritsanszky, XD2016 - A

Computer Program Package for Multipole Refinement, Topological Analysis of Charge Densities and Evaluation of Intermolecular Energies from Experimental and Theoretical Structure Factors, 2016.

[17] C. Jelsch, B. Guillot, A. Lagoutte, C. Lecomte, *J. Appl. Cryst.* **2005**, *38*, 38.

[18] B. Fournier, E. E. Bendeif, B. Guillot, A. Podjarny, C. Lecomte, C. Jelsch, *J. Am. Chem. Soc.* **2009**, *131*, 10929.

[19] a) W. Acree, J. S. Chickos, *J. Phys. Chem. Ref. Data* **2016**, *45*, 033101; b) W. Acree, J. S. Chickos, *J. Phys. Chem. Ref. Data* **2017**, *46*, 013104.

[20] a) A. Otero de la Roza, E. J. Johnson, *J. Chem. Phys.* **2012**, *137*, 054103; b) A. M. Reilly, A. Tkatchenko, *J. Chem. Phys.* **2013**, *139*, 024705.

[21] M. J. Turner, S. Grabowsky, D. Jayatilaka, M. A. Spackman, *J. Phys. Chem. Lett.* **2014**, *5*, 4249.

[22] M. J. Turner, J. J. McKinnon, S. K. Wolff, D. J. Grimwood, P. R. Spackman, D. Jayatilaka, M. A. Spackman, CrystalExplorer17, 2017. University of Western Australia. http://hirshfeldsurface.net,

[23] C. F. Mackenzie, P. R. Spackman, D. Jayatilaka, M. A. Spackman, *IUCrJ* **2017**, *4*, 575.

[24] S. P. Thomas, P. R. Spackman, D. Jayatilaka, M. A. Spackman, *J. Chem. Theory Comput.* **2018**, *14*, 1614.

[25] S. Grabowsky, T. Pfeuffer, W. Morgenroth, C. Paulmann, T. Schirmeister, P. Luger, *Org. Biomol. Chem.* **2008**, *6*, 2295.

[26] L. J. Farrugia, P. Kocovsky, H. M. Senn, S. Vyskocil, *Acta Crystallogr. B* **2009**, *65*, 757.

[27] S. Grabowsky, T. Schirmeister, C. Paulmann, T. Pfeuffer, P. Luger, *J. Org. Chem.* **2011**, *76*, 1305.

[28] M. A. Spackman, *Z. Kristallogr.* **2018**, submitted.

[29] I. Sovago, M. J. Gutmann, J. G. Hill, H. M. Senn, L. H. Thomas, C. C. Wilson, L. J. Farrugia, *Cryst. Growth Des.* **2014**, *14*, 1227.

[30] G. Rajalakshmi, V. R. Hathwar, P. Kumaradhas, *Acta Crystallogr. B* **2014**, *70*, 568.

[31] V. R. Hathwar, M. Sist, M. R. V. Jorgensen, A. H. Mamakhel, X. P. Wang, C. M. Hoffmann, K. Sugimoto, J. Overgaard, B. B. Iversen, *IUCrJ* **2015**, *2*, 563.

[32] M. V. Roux, M. Temprado, J. S. Chickos, Y. Nagano, *J. Phys. Chem. Ref. Data* **2008**, *37*, 1855.

[33] I. Sovago, M. J. Gutmann, H. M. Senn, L. H. Thomas, C. C. Wilson, L. J. Farrugia, *Acta Crystallogr. B* **2016**, *72*, 39.

[34] H. B. Bürgi, S. C. Capelli, A. E. Goeta, J. A. K. Howard, M. A. Spackman, D. S. Yufit, *Chem. Eur. J.* **2002**, *8*, 3512.

[35] A. Mukherjee, G. R. Desiraju, *Chem. Commun.* **2011**, *47*, 4090.

[36] S. R. Cox, L.-Y. Hsu, D. E. Williams, *Acta Crystallogr. A* **1981**, *37*, 293.

[37] J. Nyman, O. S. Pundyke, G. M. Day, *Phys. Chem. Chem. Phys.* **2016**, *18*, 15828.

[38] M. A. Spackman, *J. Chem. Phys.* **1986**, *85*, 6587.

[39] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787.

Poster Abstracts

in alphabetical order
(presenting authors are shown underlined)

Non-empirical distributed intermolecular force-fields for organic energetic materials

Alexander A. Aina[1], A.J. Misquitta[2], S.L. Price[1]
[1]Dept of Chemistry, Univ. College London, London, UK
[2]School of Physics and Astronomy, Queen Mary, Univ. of London, London, UK

POSTER 1 QC

An approach to deriving anisotropic atom-atom force-fields, using distributed atomic multipoles, polarizabilities and dispersion coefficients and an anisotropic atom-atom repulsion model derived from SAPT(DFT) dimer calculations, is being developed for use in predicting organic crystal structures and polymorphs. We show that this distributed intermolecular force-field (DIFF) models the experimental crystal structures of pyridine well, but raises questions about the importance of many-body terms, thermal expansion and zero-point vibrational effects.1 The DIFF model was able to identify the structure of an unreported high pressure phase of pyridine in a crystal structure prediction study, unlike an empirically fitted potential. This shows the importance of non-empirical methods for modelling the high-pressure, high-temperature phases of organic explosives. To extend this methodology to energetic materials such as TNT and RDX, the effect of the variations in the NO2 conformation on the intermolecular forces is also investigated, as the conformation can differ significantly between polymorphs. There are significant changes in the charge distribution with conformation, which are reflected in the atomic multipoles and affect the electrostatic potential around the molecule and hence the lattice energy. Proposed links between the electrostatic properties of a molecule and the observed impact sensitivities of its crystals are also investigated.

(1) Aina, A. A.; Misquitta, A. J.; Price, S. L. From dimers to the solid-state: Distributed intermolecular force-fields for pyridine. The Journal of Chemical Physics 2017.

POSTER 2 QC

In situ investigation of milling reactions and structure determination of the products using X-ray diffraction

Irina Akhmetova[1], F. Emmerling[1], K. Rademann[2], C. Roth[3]
[1]BAM Federal Inst. for Materials Research and Testing, Berlin, Germany
[2]Dept of Chemistry, Humboldt-Universität zu Berlin, Berlin, Germany
[3]Inst. for Chemistry and Biochemistry, Freie Univ. Berlin, Berlin, Germany

Mechanochemistry is a versatile approach for green and fast synthesis of pure substances. By milling the reactants, various organic, inorganic, and metal-organic compounds can be obtained in high yields. Although mechanochemistry is widely used, the underlying mechanisms are not fully understood making mechanochemical reactions difficult to predict. Metal phosphonates are metal-organic compounds accessible by grinding. Because of their structural diversity, the exploration of the chemistry of metal phosphonates has gained considerable interest during the last decades. Transition metal phosphonates are promising candidates for an application as electrocatalysts in oxygen evolution reaction (OER). Here, we present the in situ investigation of the mechanochemical synthesis of a manganese(II)-phosphonate by synchrotron X-ray diffraction and thermography. The product has not been obtained by classical solution chemistry before and its crystal structure was determined from PXRD data. The milling process can be divided into different steps, with the product crystallization corresponding with the highest temperature rise. The activity of this metal phosphonate towards OER was measured and is presented here.

Quantifying intermolecular interaction in crystals using Roby-Gould bond indices

Khidhir Alhameedi[1,2], A. Karton[1], D. Jayatilaka[1], S.P. Thomas[1]
[1]School of Molecular Sciences, Univ. of Western Australia, Crawley, WA, Australia
[2]Dept of Chemistry, College of Education for Pure Science, Univ. of Karbala, Karbala, Iraq

POSTER 3 QC

Identifying the nature of intermolecular interactions in crystals and quantifying their relative strengths is significant in the context of crystal engineering. The question whether these interactions are formed as a result of molecule·····molecule close packing or localized atom·····atom interactions is a matter of debate. In this study, we evaluate the covalent, ionic and total Roby-Gould bond index for 'σ-hole' interactions (halogen bonding and chalcogen bonding) in comparison with well-known classical hydrogen bonds. The Roby-Gould bond indices have been analyzed for a dataset of 97 crystal systems comprising 42 hydrogen bonds, 31 halogen bonding, and 33 chalcogen bonding. Our method provides insights into the nature of these interactions by separately estimating the bond indices for molecule·····molecule and atom·····atom. Hirshfeld charge for these interactions has been also reported with a trend of charge transfer from bond acceptor to the donor. A conservation law of the bond order involving the interacting atoms has been found with our Roby-Gould bond index.

POSTER 4 QC
"rising star"

HAR-ELMO - Fast and accurate Hirshfeld Atom Refinement

Lorraine Andrade Malaspina[1], A. Genoni[2], D. Jayatilaka[3], S. Grabowsky[1]
[1]Inst. of Inorganic Chemistry and Crystallography, Univ. of Bremen, Bremen, Germany
[2]Laboratory of Theoretical Physics and Chemistry, CNRS & Univ. of Lorraine, Lorraine, France
[3]School of Molecular Sciences, Univ. of Western Australia, Crawley, WA, Australia

The most widely used model nowadays for structure refinement based on X-ray diffraction data is the Independent Atom Model (IAM) due to its simplicity, speed/cost and reliability. In the IAM, atoms are considered to be spherical and non-interacting. Consequently, H atom positions (and bond lengths) derived from X-ray IAM are questionable and the obtained molecular electron density lacks chemical features.
The use of more advanced models for structure refinement which include asphericity shifts for atoms in molecules significantly improve X-H bond lengths and give access to chemically meaningful electron density, such as the multipole model [1] and Hirshfeld Atom Refinement [2,3].We will show how HAR uses tailor-made aspherical atomic scattering factors from molecular wavefunction calculations in a regular least-squares refinement procedure. However, this procedure can be very time-consuming. Only if the computational cost of HAR can be significantly reduced, its application to bigger and heavier systems will become feasible.
Extremely localized molecular orbitals (ELMOs) [4] can be used for a database-like transfer of fragment-localized orbitals, therefore expanding the regime of HAR refinement since the molecular wavefunction could be obtained almost instantaneously. The use of ELMOs in structure refinement has never been considered until now. Here, we combine HAR and ELMO and present benchmarking tests on the compounds glycyl-L-alanine and L-alanine. We compare HAR-ELMO results especially for hydrogen atom parameters with HAR results and those obtained from neutron diffraction data on the same compounds at the same temperatures.

[1] N. K. Hansen and P. Coppens, Acta Cryst. A, 34, 909-921, (1978)
[2] D. Jayatilaka and B. Dittrich, Acta Crystallogr. A, 64, 383-393, (2008).
[3] S. C. Capelli, et. al., IUCrJ, 1, 361-379, (2014).
[4] Genoni et al. – in preparation

Temperature dependence of crystal packing without phase transition in a siloxanol

POSTER 5 QC

Justin Bergmann[1], M. Fugel[1], R. Pal[1], M.F. Hesse[1], S. Mebs[2], J. Beckmann[1], P. Luger[3], S. Grabowsky[1]
[1]Inst. of Inorganic Chemistry and Crystallography ,Univ. of Bremen, Bremen, Germany
[2]Inst. of Experimental Physics, Freie Univ. Berlin, Berlin, Germany
[3]Inst. of Chemistry and Biochemistry - Inorganic Chemistry, Freie Univ. Berlin, Berlin, Germany

In this study, we investigate the crystal structures of a siloxanol with a rare intermolecular hydrogen bond.[1] The crystal packing shows significant changes upon variation of the temperature in the range from 20 K to 293 K without undergoing phase transition. In order to investigate the origin of these changes, we considered effects of the electrostatic, polarization, dispersion, and exchange-repulsion energies on the crystal packing using CrystalExplorer model energies.[2] Due to the fact that the hydrogen bond has a high impact on the intermolecular interaction energies, an accurate description of the hydrogen position is indispensable. However, the Independent Atom Model only gives an insufficient description of the hydrogen position, and we could not find any correlations between the packing energies and the hydrogen bond geometry. However, after a Hirshfeld Atom Refinement,[3] which is known to yield accurate hydrogen atom positions, such correlations were obtained.

References

[1] S. Grabowsky, M. F. Hesse, C. Paulmann, P. Luger, and J. Beckmann. How to make the ionic Si – O bond more covalent and the Si – O – Si linkage a better acceptor for hydrogen bonding. Inorg. Chem., 48(10):4384–4393, 2009.
[2] M. J. Turner, S. Grabowsky, D. Jayatilaka, M. A. Spackman: Accurate and Efficient Model Energies for Exploring Intermolecular Interactions in Molecular Crystals. J. Phys. Chem. Lett., 2014, 5, 4249
[3] S. C. Capelli, H.-B. Bürgi, B. Dittrich, S. Grabowsky, D. Jayatilaka: Hirshfeld atom refinement. IUCrJ, 2014, 1, 361-379.

POSTER 6 QC

Wet, moist or dry: Bibridged (CdCl3)- chains with a range of aqua ligand coordination

Marcus Bond, V. Buddeneni, S. Nalla
Dept of Chemistry, Southeast Missouri State Univ., Cape Girardeau, MO, USA

The [CdCl3(H2O)$^-$]n chain, an established feature of chloridocadmate structural chemistry, consists of edge-sharing, di-μ_2-chlorido CdCl5(H2O) octahedra that are are also linked by intrachain O-H…Cl hydrogen bonds that cause canting of octahedra off of the chain axis. A wealth of other hydrogen bonding opportunities—i.e. interchain O-H…O or Cl, or organoammonium cation N-H..O or Cl — are often found. We present here the "wet" chain structure found in (3-bromopyridinium)CdCl3(H2O) which is distinguished by a firmly coordinated aqua ligand and one of the longest terminal Cd-Cl bond lengths known. Multiple interchain O-H and bifurcated N-H hydrogen bonding to the terminal chloride then knit the chains together into layers. The "moist" chain structure in (3,5-dimethylpyridinium)$_2$CdCl$_3$(H_2O) contains the longest known Cd-OH_2 bond and the shortest known terminal Cd-Cl bond. In this case a short, direct N-H hydrogen bond to the aqua ligand with no hydrogen bonding to the terminal chloride contribute to this unusual bonding arrangement. Furthermore, this structure is polar ($P2_1$) by virtue of identical canting of terminal ligands on parallel chains, and suggest an application as a binary switch. Finally, the "dry" chain structure in (1,3,4-trimethylpyridindium) $CdCl_3$ does not have an aqua ligand present, and terminal chloride ligands are staggered along the chain axis to generate a "ladder pole" structure. In contrast to other "ladder pole" chain structures, in which hydrogen bonding to the terminal ligand is used to justify formation of the di-μ_2-halido bridged chain over the tri-μ_2-halido bridged chain, the quaternary cation is incapable of N-H hydrogen bonding. Computational problems of interest in these structures are (1) the energy barrier for switching polarity in the hydrate chains, (2) energetics and bonding of gradual aqua ligand removal, and (3) stability of the anhydrous di-μ_2-halido bridged chain vis-a-vis the tri-μ_2-halido bridged chain.

POSTER 7 QC

Extension of UBDB databank towards accurate electrostatic energy calculations for interactions between RNA and proteins through magnesium ions

Urszula Anna Budniak, P. M. Dominiak
Biological and Chemical Research Centre, Dept of Chemistry, Univ. of Warsaw, Warsaw, Poland

Electrostatic energy is an appropriate tool for estimating interaction energy in biomacromolecules. One of the more advanced methods to calculate electrostatic interaction energy is University at Buffalo Pseudoatom DataBank (UBDB) used together with Exact Potential Multipole Method (EPMM). Calculations are based on the structures deposited in Protein Data Bank (PDB). UBDB enables reconstruction of charge density for macromolecules in quantitative manner. By UBDB+EPMM approach, which takes also charge penetration effects into account, it is possible to compute electrostatic energies with similar accuracy as with quantum chemistry methods, for wide range of types of interactions (hydrogen bonds, π-π stacking) and distances.
The aim of my current project is to characterize electrostatic interactions in selected complexes of IFITs proteins with RNA with the use of the UBDB. IFIT5 proteins interact with pppRNA through magnesium ions. UBDB already contains many atom types including those present in amino acids and nucleotides residues, however it does not yet cover all necessary atoms of complexes containing cations e.g. Mg^{2+}. For accurate calculations, all atom types must be represented in UBDB. I would like present how new atoms are generally added and how the procedure was adapted to achieve my aim. Next I will show how to reconstruct the charge density and calculate the electrostatic interaction energy of exemplary IFIT5-pppCCCC complex. Using UBDB it is also possible to create deformation density maps and electrostatic potential maps.

Project was financed from the grant PRELUDIUM11 of National Science Centre, Poland nr 2016/21/N/ST4/03722.

Abbas YM et al. (2013) Nature 494 (7435), 60-64.
Jarzembska KN, Dominiak PM (2012) Acta Cryst. A68, 139–147.
Kumar P et al. (2014) J. Chem. Theory Comput. 10, 1652–1664.
Volkov A, Koritsánszky TS, Coppens P (2004) Chem. Phys. Lett. 391, 170–175.

POSTER 8 QC

Interaction energy calculations for hydrates of hypoxanthinium nitrate

Malgorzata K. Cabaj, P.M. Dominiak
Biological and Chemical Research Centre, Univ. of Warsaw, Warsaw, Poland

Two hydrates of hypoxanthinium nitrate were measured at different temperatures ranging from approximately 20 K to room temperature. The numerous datasets were refined with CrysAlis and Olex2. Geometry optimization, crystal energy, and theoretical structure factors were calculated using Crystal09.

The hypoxanthinium nitrate monohydrate structures exhibit interesting behavior. The first polymorph crystallizes in Pmcn group in orthorhombic crystal system, which has already been reported [1], [2]. Cooling crystal caused it to undergo a phase transition at ca. 233 K, resulting in monoclinic phase non-merohedrally twinned by two-fold axis.

The second hydrate contains a very rare occurrence of H_3O^+ cation interacting with H_2O and NO_3^-.

In our work, we wish to present an extensive comparative study of all the measured structures with analysis of the energies of interactions and geometry of the crystal packing.

[1] Rosenstein, R. D. et al., (1982). Cryst. Struct. Commun. 11, 1507-1513.
[2] Schmalle H. et al., (1990). Acta Cryst. C46, 340-342.

The influence of charge distribution on the modelling of co-crystal drug formulation

POSTER 9 QC

Joseph Cadden[1,2], S. Aitipamula[1], S.J. Coles[2]
[1]Crystallization and Formulation Science, Inst. of Chemical and Engineering Sciences, A*STAR (Agency for Science, Technology and Research), Jurong Island, Singapore
[2]Chemistry, Faculty of Natural and Environmental Sciences, Univ. of Southampton, Southampton, UK

The use of co-crystals in the pharmaceutical industry has gained considerable interest in the past decade due to their ability to improve the physiochemical properties of APIs, such as solubility, stability and bioavailability without affecting covalent bonds, or hindering pharmacological activity.[1,2]
However, the understanding of structure-property relationships, particularly within drug formulations, is limited hence studying the impact of excipients on co-crystal performance would prove valuable. To fully understand the effect of co-crystal excipient interactions upon the physiochemical properties, the combination of computational and experimental approaches has shown to be effective.
We present a co-crystal of the low-solubility API Telmisartan with saccharin, whose nature has been confirmed using PXRD, SCXRD and thermal analysis. TEL-SAC is known to exhibit significantly higher solubility than the parent API and is stable up to 6 hours *in vitro*.[3] We aim to incorporate MD simulations to model distribution of co-crystal particles within a formulation and from this predict the solubility and dissolution rate of the TEL-SAC co-crystal formulation, comparing with the experimental results. Successful prediction will also provide an opportunity to evaluate favourable cocrystal-excipient combinations.
Quantum-crystallography will also serve to complement the MD work. Currently, MD simulations provide a distribution of particles within a system, from which certain properties can be calculated, such as dissolution rate and solubility. However, these are based on atomic-resolution crystal structure data obtained from the CSD. Our target is to perform high-resolution charge-density analysis to enable calculation of a greater range of characteristics of TEL-SAC, such as co-crystal shape, density, electronic-distribution, and electrostatic potential. These can be used as improved input parameters for MD simulation calculations, thus increasing the accuracy of the model.

References
1. Thakuria, R.; Delori, A.; Jones, W.; Lipert, M. P.; Roy, L.; Rodríguez-Hornedo, N. Int. J. Pharm. 2013, 453, 101-125.
2. Blagden, N.; De Matas, M. Advanced Drug Delivery Reviews 2007, 59, 617-630.
3. Chadha et al. CrystEngComm, 2014, 16, 8375-838

POSTER 10 QC
"rising star"

A charge density study of a linear dialkyl Co(II) complex showing an unprecedented non-Aufbau electronic ground state

P.C. Bunting[1], M. Atanasov[2,3], Emil Damgaard-Møller[4], M. Perfetti[5], I. Crassee[6], M. Orlita[6,7], J. Overgaard[4], J. van Slageren[5], F. Neese[2], J.R. Long[1,8,9]

[1]Dept of Chemistry, Univ. of California, Berkeley, CA, USA
[2]Max-Planck Insit. für Kohlenforschung, Mülheim an der Ruhr, Germany
[3]Inst. of General and Inorganic Chemistry, Bulgarian Academy of Sciences, Academy Georgi Bontchev, Sofia, Bulgaria
[4]Dept of Chemistry & Centre for Materials Crystallography, Aarhus Univ., Aarhus, Denmark
[5]Inst. für Physikalische Chemie, Univ. Stuttgart, Stuttgart, Germany
[6]Laboratoire National des Champs Magnétiques Intenses, Grenoble, France
[7]Inst. of Physics, Charles Univ., Prague, Czech Republic
[8]Dept of Chemical and Biomolecular Engineering, Univ. of California, Berkeley, CA, USA
[9]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

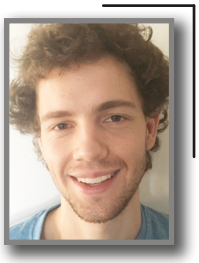
Recently the first linear dialkyl Co(II) complex has been synthesized by the Long group, on which ab initio calculations indicated a non-Aufbau electronic ground state $(d_{x2-y2}, d_{xy})^3(d_{xz}, d_{yz})^3(d_{z2})^1$. In the pursuit of confirming this abnormal electronic ground state, a variety of techniques have been used, including an experimental charge density study.

The data for the charge density analysis was obtained on the beamline BL02B1 at SPring-8, Japan using 40 keV radiation and a temperature of 20K. The resolution was sin(theta)/lamda < 1.0 and provided 8836 unique reflections with an average redundancy of 2.4 and a completeness of 97.5%.

For the charge density analysis, the Hansen-Coppens multipole formalism was used, from which a decent fit was obtained. The parameterized charge density enabled an analysis in the framework of quantum theory of atoms in molecules (QTAIM) which showed no increased covalency for the Co-C bond compared to other Co-C bonds. Given little to no covalency of the Co-C, an estimation of the 3d-orbital populations for the Co(II) ion could be obtained from the model.

The 3d-orbital analysis showed a total of 9.5 3d electrons on the Co(II) ion, which is much more than the expected 7 3d electrons, but the distribution showed 43.5% to be in the d_{x2-y2}, d_{xy} orbitals, 40.0% in the d_{xz}, d_{yz} orbitals, and 16.5% in the d_{z2} orbital. The obtained charge density thus supports the assertion of a non-aufbau ground state.

General-unrestricted extremely localised molecular orbital (gELMO) wavefunctions for Hirshfeld atom refinement (HAR)



POSTER 11 QC
"rising star"

Max Davidson, D. Jayatilaka
Dept. Molecular Sciences, Univ. of Western Australia, Crawley, WA, Australia

Currently, wavefunctions for proteins containing millions of atoms are extremely time consuming to calculate. In this poster we outline a potnetial solution to this problem and give preliminary results on small molecules for the gELMO approach. The key ideas involve the following:

(1) Use of non-orthogonal molecular orbitals, localised to a given region described by a parameter, "b". These can be localised to be atomic orbitals, allowing for a highly parallelisable procedure.

(2) Limiting interactions to another region descibed by, "p". This greatly reduces the number of required calculations and, hence, computation time.

(3) Allowing the electron spin to be chosen variationally, to compensate for the strict conditions imposed by the gELMO approximation.

With this atomic-wavefunction approach, applications to ab initio quantum crystallographic X-ray refinements, such as HAR, may be possible.

POSTER 12 QC

Towards the study of heavy-element containing solids with the CRYSTAL program: developments and applications

Jacques K. Desmarais[1,2,3], C. Ravoux[1,4], K.E. El-Kelany[1,5], R. Dovesi[1], A. Erba[1]

[1]Dip. di Chimica, Univ. di Torino, Turin, Italy
[2]Dep. of Geological Sciences, Univ. of Saskatchewan, Saskatoon, SK, Canada
[3]Dept of Physics and Engineering Physics, Univ. of Saskatchewan, Saskatoon, SK, Canada
[4]Laboratoire Structure Proprietes et Modelisation des Solides (SPMS), CentraleSupelec, Batiment G. EIFFEL, Gif-Sur-Yvette, France
[5]CompChem Lab, Chemistry Dept, Faculty of Science, Minia Univ., Minia, Egypt

The CRYSTAL [1] code for quantum-mechanical simulations of periodic systems is based on a linear-combination of atomic-orbitals (LCAO) framework. Crystalline orbitals are expressed as a linear-combination of Bloch functions that are in turn defined using a basis-set of solid-spherical Gaussian-type functions (SGTF). In its initial version, the CRYSTAL code described the wavefunction as a linear-combination of s-, p- and d- type SGTF. The code was generalized to f-type functions in 2003. Here we generalize the code to g-type SGTF, which represents the first step towards a better description of systems containing heavy-elements. Among them are lanthanides and actinides, which have occupied 4f and 5f bands and the g-type functions represent the first polarisation shell. We develop new small-core pseudopotential basis sets of the form (11sp7d8f2g)/[4sp2d3f2g] for the lanthanide and actinide series in the solid state.

The new code and basis sets are applied to the strongly-correlated lanthanide sesquioxides Ln_2O_3 series (Ln=La,Ce,Pr,Nd). We provide a detailed description of their geometric, electronic and magnetic structures using global-hybrid functionals. The critical role of the fraction of Fock exchange is addressed. In particular, a new theoretical approach based on a self-consistent definition - through the material's dielectric response - of the optimal fraction of exchange in hybrid functionals [2] is applied for the frist time to strongly correlated materials. Our description of Ce_2O_3 is in agreement with the existing copious litterature on this system, which suggest that our description of the other members of the series is reliable, for which very little has been published before.

[1]Dovesi R, Erba A, Orlando R, Zicovich-Wilson CM, Civalleri B,Maschio L, Rérat M, Casassa S, Baima J, Salustro S, Kirtman B (2018) WIREs Comput Mol Sci
[2] A. Erba, J. Phys.: Condens. Matter 29, 314001 (2017).

Wavefunctions fitted to charge density data – exploring the effects of data quality and obtainable properties

Jonathan J. Du[1], A. Genoni[2], J. Overgaard[3], D.E. Hibbs[1]
[1]Faculty of Pharmacy, The Univ. of Sydney, NSW, Australia
[2]CNRS & Univ. de Lorraine, Laboratorie LPCT, Metz, France
[3]Dept of Chemistry, Center for Materials Crystallography, Aarhus Univ., Aarhus, Denmark

POSTER 13 QC
"rising star"

The fitting of a wavefunction to experimental diffraction data has long been considered as an alternative to the current multipole model in refining the charge density of crystal systems. Its primary advantage lies in its ability to obtain the density matrix and, above all, density matrix-related properties (such as the kinetic energy, Mayer bond orders and Weinhold population analysis) that the multipole model can only estimate. Other known issues associated with the multipole model, including the correlation between anisotropic displacement parameters (ADP's) and multipole parameters and the presence of regions of negative density are also addressed using wavefunction-based methods.

In this study, charge density data for the anti-epileptic drug carbamazepine were collected at the Faculty of Pharmacy at the University of Sydney at 150K and at the SPring-8 synchrotron facility at 20K. Three wavefunction fitting methods (Jayatilaka's approach implemented in TONTO[1], the XC-ELMO strategy by Genoni[2] and the Molecular Orbital Occupation Number (MOON) refinement techniques by Hibbs *et al.*[3]) were used to generate wavefunctions fitted to experimental diffraction data for both the two data sets. This investigation examines the effect of data quality and choice of basis set used for the fitting process on the final obtained wavefunctions and compares the properties obtained from each method. The work aims at providing a concise summary of the benefits offered by each fitting method and a guide as to which method is best if a particular property of the charge density is desired. Further studies will be aimed towards finding the optimal functional and basis sets to be used for a variety of compounds including systems with Z' >1 such as pharmaceutical co-crystals.

POSTER 14 QC
"rising star"

Bonding in polyiodides

Michelle Ernst, T. Poreba, P. Macchi
Dept of Chemistry and Biochemistry, Univ. of Bern, Bern, Switzerland

We are studying the bond formation in polyiodides induced by applying pressure on a crystal. Besides the analysis of the structure obtained from X-ray diffraction we use various theoretical tools such as bond decomposition analyses and X-ray constrained wavefunctions (XCW). While traditional methods for the refinement of X-ray diffraction data, like for example the multipolar refinement, provide only information about the diagonal elements of the one-electron reduced density matrix (the electron density), the XCW method enables the calculation of the density matrices as well, although the physical meaning is not fully understood. XCWs are obtained by calculating a wavefunction which simultaneously minimizes the energy obtained as expectation value of the Hamilton operator and an agreement statistics obtained by comparing the calculated structure factors to the measured ones.

Polyiodides form a very versatile class of compounds. Our model crystal consists of three iodine units (I_2, I_3^- and I_2) forming a zig-zag chain. Studying this crystal under high pressure allows the observation of structural changes and bond formation within the very same crystal.
X-ray diffraction measurements were done at different pressure points up to 12 GPa. At ambient pressure the distance of the two I_2 units to the I_3^- unit is equivalent (and restricted by symmetry), around 6 GPa the crystal undergoes a phase transition which results in the formation of an I_5^- and upon further compression an I_7^- unit. However, no unique criteria for what is to be considered a bond exist. We used the EDA and IQA approaches together with other criteria such as the delocalization index and the Laplacian of the electron density to clarify the bonding situation. Additionally, orbitals obtained from XCW calculations were analysed in order to better understand the interactions that lead to the bond formation.
In future we plan to use X-ray constrained ELMOs to localize orbitals on the interacting fragments.

Quantum Crystallography put into action: assessing the accuracy of and getting insights on the electron spin densities

Carlo Gatti[1,2], G. Macetti[3]
[1]CNR-ISTM, Ist. di Scienze e Tecnologie Molecolari, Milan, Italy
[2]Ist. Lombardo Accademia di Scienze e Lettere, Milan, Italy
[3]Dip. di Chimica, Univ. degli Studi di Milano, Milan, Italy

POSTER 15 QC

Modelling magnetic properties of crystalline compounds by experiments is truly challenging.[1] New methodological routes are now available to ease this task. By combining the information obtained by different techniques such as X-ray and polarized neutron diffraction[2,3] or polarized neutron and X-ray magnetic diffraction,[4] the limitation of the scarce number and resolution of experimental data is partly overcome. More precise wavefunction-based models or electron spin densities (SDs) are so obtained and comparison with *exact* wavefunctions or densities becomes clearly advisable to test their accuracy.
Ab-initio electron SDs can be easily calculated but they are usually far from being reliable. Systematic studies demonstrated that DFT is often unable to treat open-shell systems properly, leading to non-accurate electron SDs.[5,6] *Ab-initio* electron correlation methods or density-matrix renormalization group approaches,[6] have to be called for. Yet, they are computationally too demanding and unsuited for large systems. Extraction of chemical information from the SD is not a trivial task as for the ED, even for very simple molecular systems.[7] The concept of Source Function (SF)[8] was thus purposely extended to SD distributions and applied to analyze the magnetic patterns in metal complexes molecular crystals.[9,10] The SF SD serves as a useful tool for discussing the SD accuracy and to disclose the origin of the SD discrepancies when approaches of increasing quality are used.

[1] A. Genoni et al., Chem. Eur. J. 10.1002/chem.201705952
[2] M. Deutsch, et al., IUCrJ 2014, 1, 194
[3] M. Deutsch et al., Acta Cryst. A 2012, 68, 675
[4] I.A. Kibalin et al., Phys. Rev. B 2017, 96, 054426
[5] K. Boguslawski et al., JCTC 2011, 7, 2740
[6] K. Boguslawski et al., JCTC. 2012, 8, 1970
[7] C. Gatti et al.,Chem. Sci. 2015, 6, 3845
[8] R.F.W. Bader & C. Gatti, Chem. Phys. Lett. 1998, 287, 233
[9] G. Macetti et al., J. Comp. Chem. 2018, 39, 587
[10] C. Gatti et al., Acta Cryst B 2017, 73, 565

POSTER 16 QC

Structural analysis of Bornite in the range 10 K-275 K: a synchrotron light and DFT investigation

Andrea Giaccherini[1,2], A. Martinelli[3], G.O. Lepore[4], F. Bernardini[5], F. Di Benedetto[6]
[1]Dept of Earth Sciences, Univ. of Florence, Florence, Italy
[2]Dept of Chemistry, Univ. of Florence, Sesto Fiorentino, Italy
[3]SPIN-CNR, Genoa, Italy
[4]IOM-CNR, European Synchrotron Radiation Facility 71, Grenoble, France
[5]Dept of Physics, Univ. of Cagliari, Monserrato, Italy

Bornite is a widespread sulfide mineral with nominal composition Cu_5FeS_4 perfectly fitting the appealing characteristics of technologically relevant multinary sulphides for thermoelectric (TE) technology. Bornite has a recent history of success due to its interesting properties, non-toxicity and earth-abundance of its constituting elements. During the last 40 years, its structure was the subject of several investigations at different temperatures with no conclusive structural determination of its polymorphs. In particular, several works suggest a structural ordering of Fe(III) in two specific sites at room temperature, confirmation of these observation requires further investigation. To this aim, we improved the quality of the diffraction data available and attempted a further structural determination at cryogenic temperature. Specifically, we investigated a natural Bornite sample by means of high resolution synchrotron X-ray powder diffraction (XRD), pair distribution function (PDF) analysis and X-ray absorption spectroscopy (XAS) between 10 K and 275 K. The experimental data confirmed the Pbca space group and strongly supports the preferred location of Fe in bornite at 275 K. Regarding the changes at cryogenic temperature, we found that the unit cell volume decreases continuously with decreasing temperature, undergoing an abrupt contraction below ~65 K, where a 1st order Pbca → $Pca2_1$ structural transition takes place. The analysis of the vibrational mode yielded the primary active mode breaking the Pbca symmetry towards $Pca2_1$. The Pbca → $Pca2_1$ structural transition was further investigated by first-principles calculations at DFT-PBE level of theory. Starting from the optimized structure only three phonon mode frequencies have been found lower than 5.6 THz with a finite imaginary component of about 6 THz. These modes distort the structure from Pbca to $Pca2_1$ confirming the experimental findings.

The riddle of helium arsenolite inclusion compound formation

Piotr Guńka[1], M. Hapka[2], M. Hanfland[3], M. Dranka[1], G. Chałasiński[2], J. Zachara[1]
[1]Faculty of Chemistry, Warsaw Univ. of Technology, Warsaw, Poland
[2]Faculty of Chemistry, Univ. of Warsaw, Warsaw, Poland
[3]European Synchrotron Radiation Facility, Grenoble, France

POSTER 17 QC

Investigations into the helium permeation of arsenolite, the cubic, molecular arsenic(III) oxide polymorph As_4O_6, were carried out to understand how and why arsenolite helium clathrate $As_4O_6{\cdot}2He$ is formed. High-pressure synchrotron X-ray diffraction experiments on arsenolite single crystals revealed that the permeation of helium into nonporous arsenolite depends on the time for which the crystal is subjected to high pressure and on the crystal history. The single crystal was completely transformed into $As_4O_6{\cdot}2He$ within 45 h under 5 GPa. After releasing the pressure, arsenolite was recovered and a repeated increase in pressure up to 3 GPa led to practically instant $As_4O_6{\cdot}2He$ formation. However, when a pristine arsenolite single crystal was quickly subjected to a pressure of 13 GPa, no helium permeation was observed at all. No neon permeation was observed in analogous experiments. Quantum mechanical computations allow for the modelling of inclusion compound formation reaction enthalpy and indicate that there are no specific attractive interactions between He atoms and As_4O_6 molecules at the distances observed in the $As_4O_6{\cdot}2He$ crystal structure. Detailed analysis of As_4O_6 molecular structure changes has shown that the introduction of He into the arsenolite crystal lattice significantly reduces molecular deformations by decreasing the anisotropy of stress exerted on the As_4O_6 molecules. This effect and the $p\Delta V$ term, rather than any specific As···He binding, are the driving forces for the formation $As_4O_6{\cdot}2He$.

POSTER 18 QC

Expanding crystal structure prediction to larger and more flexible molecules of pharmaceutical interest

Luca Iuzzolino
Dept of Chemistry, Univ. College London, London, UK

Crystal Structure Prediction (CSP) studies aim to predict all the thermodynamically plausible crystal structures of a molecule from the chemical diagram, and can be used in the pharmaceutical industry as a complement to the solid form screening work carried out in drug development. The industrial use of CSP is currently limited by the huge computational cost, which scales very badly with the size and flexibility of the molecule. Hence, we investigate methods to reduce the computational cost of CSP studies of large flexible molecules. Conformational information retrieved from the Cambridge Structural Database (CSD), including using the CSD Conformer Generator, can facilitate the exploration of the lattice energy (E_{latt}) surface for generating millions of plausible crystal structures. Dispersion-corrected tight-binding density functional theory (DFTB-D) can be used as an intermediate step to relax all inter and intra-molecular degrees of freedom of several thousands of approximate crystal structures generated in a search. This reduces the number and cost of the energy calculations that need to be performed with higher levels of accuracy. These methods are tested for five large flexible molecules by comparison with previous extensive CSP studies. The molecules include the anthelmintic drug mebendazole, where the original CSP study was recently performed to complement and guide an extensive polymorph screen. The revised methodology successfully generated and gave a good energy ranking to almost all the low-energy crystal structures found in the original CSP studies, including all experimental structures. This was achieved at significantly reduced computational cost compared with the original studies, and an analysis of the results of the tests suggests that the approach can be extended to larger and more flexible molecules.

QM/MM-based structure and charge density analysis of Biomolecules: A new approach for drug design

Saravanan Kandasamy, K. Poomani
Laboratory of Biocrystallography and Computational Molecular Biology,
Dept of Physics, Periyar Univ., Salem, India

POSTER 19 QC

The favorable outcome of drug design depends on the intermolecular interaction between the target protein and ligand molecules. Here, the geometry of these interactions is being used to decide the strong and weak intermolecular interactions. Nowadays, the third generation high-throughput advanced crystallography techniques allow to investigate these interactions at electron density level using high resolution of X-ray diffraction. This electronic level information gives deeper understanding the structural facts with the position of hydrogen atoms and valence electron distributions; in particular, understanding the electronic properties of drug-receptor interactions is very much essential to elucidate the molecular mechanism. However, the high resolution of X-ray diffraction is becoming very common in small molecular charge density studies and it is not straightforward in protein-ligand complexes due to limited X-ray diffraction data and a large number of parameters. To make it very feasible, in the present study we adopted a QM/MM computational method, in which we considered the drug molecule and the interacting amino acids in the active site (binding pocket) as a QM region and the remaining part of the complex as an MM region. Since this model has a physiological relevance, we proceed to determine the structure, charge density distribution and topological properties of the intermolecular interaction of biomolecules in the active site of the proteins. This theoretical onsite charge density analysis of drug-receptor complex reveals the nature of charge density distribution of binding pocket using QM/MM method coupled with QTAIM analysis. Here, few drug molecules with corresponding enzymes against Alzheimer disease were investigated and the results will be discussed at the time presentation.

POSTER 20 QC

The Quantum Crystallography of Sila-Ibuprofen and its application

Florian Kleemiss[1], M. Fugel[1], K. Sugimoto[2], J. Beckmann[1], S. Grabowsky[1]
[1]Univ. of Bremen, Bremen, Germany
[2]JASRI, SPRing-8, Hyogo Prefecture, Japan

We have synthesized the compound sila-ibuprofen and shown its promise in biochemical tests. We characterized sila-ibuprofen in comparison to the most common anti-inflammatory painkiller, ibuprofen. To get an insight into the differences and similarities between common ibuprofen and sila-ibuprofen a series of quantum crystallographic analyses were performed: X-ray wavefunction refinement (XWR) [1] was carried out on high-resolution low-temperature diffraction data obtained at BL02-B1 of SPring-8 in Hyogo, Japan. It was followed by model energy analysis of the crystal packing [2] and Non-Covalent-Interaction(NCI-)-plots. [3] This is necessary since classical structural refinements are neither capable of locating hydrogen atoms accurately, nor describing bonding effects such as electrostatic potentials. The effect of the crystal environment of a small biologically active ligand in the crystal packing of the pure substance on its electron density has been claimed to be a useful approximation of the polarization in the biological situation. [4]
The structure and wavefunction obtained through XWR refinement were used to develop a force field to simulate the interactions of the substance in the active site of the enzyme using molecular dynamics. The results of these investigations show a strong correlation of properties observed in the crystal and obtained from simulations of the enzyme pocket. It might therefore explain the biochemically observable trends of activity exhibited by sila-ibuprofen.

[1] M. Woinska, D. Jayatilaka, B. Dittrich, R. Flaig, P. Luger, K. Wozniak, P. M. Dominiak, S. Grabowsky, Chem. Phys. Chem. 2017, 18, 3334–3351.
[2] M. J. Turner, S. Grabowsky, D. Jayatilaka, M. A. Spackman, J. Phys. Chem. Lett. 2014, 5, 4249.
[3] J. Contreras-Garcia, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D.N. Beratan, W. Yang, J. Chem. Theory Comput. 2011, 7 (3), 625–632.
[4] C. Pascard, Acta Cryst. D 1995, 51, 407-417.

Influence of chosen synthons on the polarizabilities of functional groups

Anna Krawczuk
Faculty of Chemistry, Jagiellonian Univ., Krakow, Poland

POSTER 21 QC
"rising star"

The correlation between the crystal structure and physical properties of a given material has long been a subject of many studies. One of the key features of designing efficient multifunctional materials is to use specific building blocks and/or synthons in order to increase a desired effect in the certain crystallographic direction. For example, to obtain efficient optical devices it is necessary to use highly polarizable functional groups which will promote high optical effect. It is thus crucial to get a precise information on how those group polarizabilities are influenced by common synthons used in crystal engineering.
Recently developed routine, PolaBer[1] allows us to calculate atomic polarizabilities, and therefore group polarizabilities, based on the definition of atomic dipole moments first given by Bader[2] and developed by Keith.[3] The routine uses the results of QTAIM partitioning of charge densities. According to QTAIM theory each atomic contribution can be expressed as a sum of *atomic polarization* $\boldsymbol{m}_p$(W) and *charge translation* m_c(W) terms. The numerical derivatives of these quantities in respect to external electric field provide the atomic polarizabilities. The advantage of using atomic polarizabilities rather than molecular ones, is the fact that we can extract separate information about the atomic and intermolecular contribution into linear susceptibility. This, on the other hand, enables us to identify which group mostly contributes to the global dielectric constant, which in turn could be very helpful for the crystal engineering purpose when designing new optically effective materials.

References
[1] A. Krawczuk, D. Pérez, P. Macchi J. Appl. Cryst. 47 (2014) 1452-1458.
[2] R.F.W. Bader, Atoms in Molecules: A Quantum Theory, Oxford University Press, Oxford, U.K., 1990.
[3] Keith, T.A. (2007) Atomic Response Properties in The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design Matta C.F., Boyd R.J., Eds., Wiley-VCH, Weinheim, 2007.

POSTER 22 QC

Combined spectroscopic and crystallographic studies of new photoswitchable Ni and Cu-based nitrosystems in the solid state

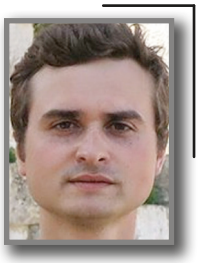
Sylwia E. Kutyla[1], A. Krówczyński[1], R. Kamiński[1], D. Schaniel[2], K.N. Jarzembska[1]
[1]Dept of Chemistry, Univ. of Warsaw, Warsaw, Poland
[2]CRM2, UMR 7036, Univ. de Lorraine, Vandoeuvre-les-Nancy, France

Stimuli-responsive chemical systems exhibiting specific photoactive properties have gained a lot of attention nowadays due to their potential technological applications (solar cells, LEDs, biological markers, ect.). It is, thus, of great importance to understand the phenomena behind the properties of interest, so as to design the desired materials and sensibly control their properties.
Hence, the presented project has been dedicated to synthesis and crystallization of novel Ni and Cu nitro complexes, as well as their further detailed characteristic using X–ray diffraction and spectroscopic techniques. The studied systems were examined specifically for their ability to undergo the nitro group isomerisation. Conversion between different isomers can be achieved in two ways, thermally and/or via photo-activation. In the ground state the nitro group exhibits the nitro binding mode (η1-NO2) in the case of nickel complexes, whereas the *nitrito* (η1-ONO) linkage in the copper analogues.
So far, the most promising complex from the examined series, *i.e.* PhTNiNO_2, (PhT = 2 - [8-quinolyloamino) methyl] 1-tetralone), exhibits full conversion when irradiated with the LED diode 590 nm (green light) or 660 nm (red light) at 160 K for 1 hour. Full conversion was earlier reported by Warren *et al.* [1], for only one Ni complex. It should, however, be stressed that the newly synthesized compound constitutes the first Ni-complex for which the metastable form is stable up to 240 K (for comparison, the metastable state of Ni(dppe)(NO_2)Cl is stable up to 150 K). Other nickel complexes analysed up to date exhibit comparable, but slightly worse properties. In turn, the copper systems work best at 10 K, whereas the metastable form is usually stable up to 50-80 K, which makes them more difficult to be analysed and less applicable materials.

Financial support from the PRELUDIUM grant (2017/25/N/ST4/02440) of the NCN in Poland is gratefully acknowledged.

[1] M.Warren, J. Chem. Eur. 2014,20,5468-5477.

Experimental observation of Jahn-Teller distortions in π-conjugated high symmetry systems: C_{60}^{n-} and $[MPc]^{n-}$ anions

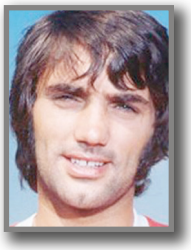


POSTER 23 QC
"rising star"

Alexey Kuzmin[1], A. Krylova[1], M. Faraonov[2], A. Fatalov[2]
[1]Inst. of Solid State Physics of Russian Academy of Sciences, Chernogolovka, Russian Federation
[2]Inst. of Problems of Chemical Physics of Russian Academy of Sciences, Chernogolovka, Russian Federation

The theory of Jahn-Teller (J-T) effect has concluded that a degenerate electronic state of a high symmetry system is unstable to any deformations. The reduced forms of fullerene (C_{60}) and metal phthalocyanine (MPc) are subject to J-T distortions due to extra electrons on the degenerated LUMO levels. The distortion removes the degeneracy of this orbital and splits it into several levels with slightly different energies.
The original J-T theorem does not say which distortions is the most preferable and does not give any estimation of their magnitudes. Ab initio and DFT calculations show that low symmetry conformers are placed very close to each other on a potential energy surface, so the question regarding which of them is realized in real crystals is complicated.
The vast majority of works on this issue contains the theoretical study of isolated ions with optimized geometry; however, it takes place only in a gas phase. In order to find the experimental evidence of such effects, we have analyzed the structural data for the bunch of molecular and ionic complexes based on fullerenes and phthalocyanines. Most of the structures were obtained in our group during last four years. A complete geometry analysis of the molecular cage of C_{60}^{n-} and $[MPc]^{n-}$ depending on their negative charge n was proposed. It has been found that the Pc reduction affects the molecular structure of metal phthalocyanine radical anions and leads to the alternation of short and long C-N bonds in the heterocycle and destroys D_{4h} symmetry of the molecule. In the case of C_{60}^{n-} the molecular shape was approximated by the principal ellipsoid, and the distortion was analyzed by comparing the axes lengths. The way of LUMO splitting depends on distortion parameters. EPR and UV-vis spectroscopy was used to confirm the results based on X-ray atomic structure analysis.

This work was partially supported by RFBR according to the research project № 18-33-00731.



Synthesis and crystal structure of new alkaline-chalcogenido-manganates

Michael Langenmaier, C. Röhr
Inst. für Anorganische und Analytische Chemie, Albert-Ludwigs-Universität, Freiburg, Germany

POSTER 24 QC

Until now alkali-rich chalcogenido-metallates A_6MQ_4 [1-5] (M=Mn/Fe/Co/Zn; Q=S/Se/Te) are only known for the lighter alkali elements Na and K. These ortho salts with isolated $[TMQ_4]^{6-}$ tetrahedra crystallize in the Na_6ZnO_4-type structure [5].
The new compounds Rb_6MnS_4, Rb_6MnSe_4, Rb_6MnTe_4, Cs_6MnS_4, Cs_6MnSe_4 and Cs_6MnTe_4 were obtained during the systematic studies on the crystal chemistry of chalcogenido manganates [6].
All compounds are isotypic and crystallize in the hexagonal spacegroup $P6_3mc$, with lattice parameters from a=1021.9(7) and c=796.2(5) pm (Rb_6MnS_4) to a=1173.3(7) and c=906.3(5) pm (Cs_6MnTe_4). The crystal structure, which was determined by means of X-ray single crystal diffraction, contains two A, one Mn and two Q sites. The Q atoms form a hexagonal close-packing with slightly undulated hexagonal layers due to Q(1) being slightly out of the plane.The Mn(II) ions are located in nearly ideal Q-tetrahedra and occupy 1/8 of the vacancies. These isolated tetrahedra are the key characteristic of the structure and are uniformly oriented along [001]. The A(1) ions take 3/8 of tetrahedral intersticies and share three common edges with the aforementioned $[MnQ_4]$ tetrahedra.
The second type of A^+ ions are octahedrally coordinated by six Q^{2-} anions. These octahedra share common edges with themselves as well as common faces with both kinds of tetrahedra. The remaining 1/4 of octahedra voids remain empty and are connected via common faces, thus forming channels along the c axis.

[1] K. Klepp, W. Bronger, Rev. Chim. Miner., 20, 682 (1983).
[2] K. Klepp, W. Bronger, Z. Naturforsch. 38b, 12 (1983).
[3] W. Bronger, H. Balk-Hardtdegen, Z. Anorg. Allg. Chem., 574, 89 (1989).
[4] W. Bronger, H. Balk-Hardtdegen, U. Ruschewitz, Z. Anorg. Allg. Chem., 616,14 (1992).
[5] P. Kastner, R. Hoppe, Z. Anorg. Allg. Chem., 409, 69 (1974).
[6] M. Langenmaier, C. Röhr, Z. Kristallogr. Suppl. 37, 113 (2017).

Quantitative and qualitative analysis of hydrogen bonds and interaction energies in proteins

Suman Kumar Mandal[1], B. Guillot[2], P. Munshi[1]
[1]Chemical and Biological Crystallography Lab, Dept of Chemistry, Shiv Nadar Univ., Gautam Buddha Nagar, Uttar Pradesh, India
[2]Laboratoire de Cristallographie, Inst. Jean Barriol, Univ. de Lorraine, Nancy, France

POSTER 25 QC

The function of a protein is defined by its structure and it is well perceived that the folding of proteins to its 3D structure occurs due to the burial of hydrophobic side chains and the formation of main-chain N-H•••O=C hydrogen bonds (HB). Although there is a debate on the predominating factor for the folding process, recent studies direct towards the electrostatic HB interactions.[1] In globular proteins, N-H•••O=C contacts, formed between the main chain N-H and C=O groups, are the building block of α-helices and β-sheets. It is expected that the accurate information of such interactions in proteins will help in proper modeling and prediction of binding energies. Given the recent developments in macromolecular crystallography and with the rise of quantum crystallography concept,[2] it is encouraging to study electrostatic HB interactions using experimentally derived electron densities. Research on protein charge density analysis has recently been stimulated particularly with the interest in the topological analysis of HBs, protein-ligand interactions and their electrostatics.[3] In this study, we have modeled charge densities (ELMAM-2[4] based), of several high-resolution X-ray protein structures as deposited in the PDB and performed topological analysis followed by electrostatic interaction energy estimation. Further, the topologies of the HBs are visualized[5] and the interaction energies (E) are quantified. The topological properties thus obtained are found to correlate well among themselves as well as with the E. Finally, this systematic study demonstrated that the N-H•••O=C hydrogen bonds in proteins follow the same trend as those studied in case of small molecules.[6]

1. Rose et.al., 2006 PNAS, 103, 16623
2. Grabowsky et.al., 2017 Chem Sci, 8, 4159
3. Liebschner et.al., 2011 J Phy Chem. A, 115, 12895
4. Domagala et.al., 2012 Acta Cryst A, 68, 337
5. Saleh et.al., 2013 J Appl Cryst, 46, 1513
6. Mallinson et.al. 2003 JACS, 125, 4259; Munshi & Row 2006 Cryst Eng Comm, 7, 608

POSTER 26 QC

Theoretical and experimental studies of conformational changes induced by intermolecular interactions in various crystals of 5-benzofurazancarboxylic acid

Paulina H. Marek[1,2], I. Madura[1]
[1]Faculty of Chemistry, Warsaw Univ. of Technology, Warsaw, Poland
[2]The Faculty Laboratory of Advanced Crystal Engineering, Faculty of Chemistry, Univ. of Warsaw, Warsaw, Poland

One of the computational ways to support crystal engineering are crystal structure prediction methods. It can be particularly useful in predicting structures of labile molecules. However, these methods are extremely time-consuming, therefore, simple methods allowing initial estimation of the most energy preferable conformations of the molecule are being sought. The aim of the research was to propose a simplified approach of assessing the most preferred angular ranges of spatial alignment of carboxylic groups in dicarboxylic acids. Performed predictions were followed by comparison of the results with the data available in the CSD crystallographic database. The method confirmed that rough calculations conducted in gas phase could be used to determine energetically preferred molecular geometries. What more, it was concluded that by modification of chemical environment of a given molecule in crystal, one can influence its conformation. In the next step, proposed method was used to estimate energy values in molecules of asymmetric monocarboxylic acids. The results indicated that it would be theoretically possible to observe both acid conformers in the crystalline state. Basing on the calculation results, a number of multi-component systems with 5-benzofurazancarboxylic acid was designed. The role of the second component was to modify the chemical environment of the acid molecule, thereby inducing conformational pseudopolymorphism. Structures were examined by X-ray diffraction methods. Hirshfeld surface analysis along with graph theory were used to describe intermolecular interactions of obtained structures. Results of this studies can be used as a starting point for analysis of the conformational polymorphism phenomenon in the structures of asymmetric monocarboxylic acids.

Unprecedent stepwise investigation of guest loading in highly flexible MOF pores by crystalline sponge method

Paolo P. Mazzeo[1,2], D. Balestri[1], C. Carraro[1], P. Pelagatti[1], N. Demitri[3], A. Bacchi[1,2]
[1]Univ. of Parma, Parma, Italy
[2]Centro Interdipartimentale Biopharmanet-TEC, Parma, Italy
[3]Sincrotrone Elettra, Trieste, Italy

POSTER 27 QC

MOFs are known to be highly versatile materials made by connecting metallic nodes characterized by a given coordination geometry with multidentate rigid ligands acting as spacers, hence affording three-dimensional network connectivity. The accurate design of the building units allows to obtain porous MOFs with cavities of considerable size which usually accommodate loosely bound solvent molecules. Starting from the pioneer concept recently developed of crystalline-sponge method[1] in which the authors focused their attention on the structural determination of single guest molecules trapped into a microporous framework, we here propose a systematic way to embed small molecular aggregates inside home-made porous crystalline materials, with the multiple aims of exploring both the structural aspects of nanoconfinement and the stabilization of guest molecules inside the cavities of the structure. The challenge of this idea stands in the possibility to neatly "freeze", within a crystal, ordered supramolecular clusters of molecules that would form a liquid in their natural state at ambient conditions, and visualize their supramolecular aggregation, at different loading time. In particular, we here propose an unprecedent stepwise characterization of the evolution of the guest loading process by soaking the MOF crystals for different time into the pure liquid guest and sequentially investigate them via SCXRD. The guest we focused on is Eugenol, which is a volatile phenolic constituent of clove essential oil obtained from Eugenia caryophyllata buds and leaves. It is a functional ingredient of numerous products which have been used in the pharmaceutical, agro-food and cosmetic industry. The wide range of eugenol activities derived from it antimicrobial, anti-inflammatory, analgesic and antioxidant properties.

Switching from quasi-2D to 3D Heisenberg antiferromagnets in $\{[Cu(pyz)_2X]BF_4\}_n$ with X=Cl and Br

Fabio Montisci[1], R. Scatena[1], A. Lanza[1,2], N. Casati[3], P. Macchi[1]
[1]Dept for Chemistry and Biochemistry, Univ. of Bern, Switzerland
[2]Center for Nanotechnology Innovation, Italian Inst. of Technology, Pisa, Italy
[3]Swiss Light Source, Paul Scherrer Inst., Villigen, Switzerland

POSTER 28 QC

The design and control of materials with quasi-low dimensional antiferromagnetic Heisenberg interactions is desirable for relevant applications in spintronics. $\{[Cu(pyz)_2X]BF_4\}n$ are excellent models to investigate low dimensional systems because of their localized spin-1/2 moments and large charge gap. We studied these samples both experimentally and with periodic DFT calculations, determining the electron density distribution through high resolution single crystal X-ray diffraction and probing structural and electronic changes trough application of high pressure.
The analysis of the electron and spin density distribution, as well as their correlation with magnetic properties, was employed to describe the super-exchange mechanism. Moreover, models using restricted multipoles were refined against the high pressure X-ray diffraction data and analyzed with QTAIM.
The species shows remarkable stability of the magnetic network up to quite high pressure (above 10 GPa). The experimentally observed antiferromagnetic coupling can be explained by the super-exchange coupling mediated by pyrazine linkers. However, the exchange through pyrazine results to be modulated by the interaction with the apical linker X, which at increased pressures acquires relevance and transforms the magnetic network from quasi-2D to 3D.
The results of this study suggest that systematic electron density analysis on transition metal compounds could lead to a better understanding of the super-exchange mechanism.

Structural characterization and computational study of spryite, $Ag8As^{3+}_{0.5}As^{5+}_{0.5}S_6$

POSTER 29 QC

Marta Morana[1], L. Bindi[2], F.N. Keutsch[3], F. Zaccarini[4], G. Cardini[5]

[1]Dept Earth and Einviromental Sciences, Pavia Univ., Pavia, Italy
[2]Dept Earth Sciences, Florence Univ., Florence, Italy
[3]Dept Chemistry and Chemical Biology, Harvard Univ., Cambridge, MA, USA
[4]Dept Applied Geosciences and Geophysics, Leoben Univ., Leoben, Austria
[5]Dept Chemistry Ugo Schiff, Florence Univ., Florence, Italy

Natural and synthetic Ag-sulfosalts are interesting compounds because of their distinctive physical and chemical properties, such as ionic conductivity. Sulfosalts are characterized by some of the most complicated atomic and crystal structures known in themineral kingdom. Among them, argyrodites ($A^{m+}_{(12-n-y)/m}B^{n+}X^{2-}_{6-y}Z^{-}_{y}$) show a complex behaviour as a function of temperature, and superionic conductivity is often found to be associated with the high-temperature modification. The high mobility of the monovalent cations (i.e., Cu and Ag) is often responsible of the strong structural disorder of the respective ions. We recently determined the crystal structure of a previously unknown natural Ag-bearing sulfosalt, $Ag^8(As^{3+}_{0.5}As^{5+}_{0.5})S_6$, which was named spryite (Bindi et al.2017). The mineral exhibits a crystallographic peculiarity: the position usually occupied by Ge in argyrodite corresponds to two partially-occupied sites in spryite. The first position has a tetrahedral environment, while the second has a trigonal pyramidal coordination. Consequently, As^{5+} has been thought mixed with Ge^{4+} at the first position, whereas As^{3+} to be hosted in the other position. Thus, As^{3+} and As^{5+} coexist in spryite, that represents the first As^{3+}-bearing member of the argyrodite group. The presence of the partially occupied $As^{3+}S_3$ pyramids seems to inhibit the formation of the "quasi-liquid-like"structure that is usually associated to high temperature polymorphs and fast ion conductors. Accordingly, no phase transitions were detected in the high-temperature X-ray diffraction experiments. Density functional theory calculations were performed to obtain more information about the electronic structure.

References

Bindi, Luca and Keutsch, Frank N and Morana, Marta and Zaccarini, Federica Zaccarini (2017) Spryite,$Ag^8(As^{3+}_{0.5}As^{5+}_{0.5})S_6$:structure determination and inferred absence of superionic conduction of the first As3+-bearing argyrodite, Physics and Chemistry of Minerals

POSTER 30 QC

Hirshfeld atom refinement for an organo-gold(I) compound

Sylwia Pawlędzio[1], M. Woińska[2], A. Makal[1], K. Woźniak[1]
[1]Biological and Chemical Research Centre, Dept of Chemistry, Univ. of Warsaw, Warsaw, Poland
[2]Dept of Molecular Physiology and Biological Physics, Univ. of Virginia, Charlottesville, VA, USA

Relativistic effects in chemistry manifest themselves in many ways and influence various physical and chemical properties of materials. The well-known macroscopic examples are: the yellow colour of gold[1] and the low melting point of mercury[1]. The relativistic effects appear when the speed of electrons approaches the speed of light[1] and for valence shell they increase like Z^2. They manifest themselves in three major aspects[2]: (1) the radial contraction and decrease of energy of the *s* and *p* shells, (2) the spin-orbit splitting and the radial expansion and (3) increase of the energy of upper *d* and *f* shells. Single crystal X-ray diffraction provides information about the electron distribution in molecules, which can be modelled with the Hansen-Coppens formalism[3](MM). Unfortunately, the MM is hardly flexible enough for heavy atoms. An alternative method is the X-ray Wavefunction Refinement (XWR)[4] which consists of two steps: Hirshfeld Atom Refinement (HAR) and XCW fitting. According to Bucinsky *et al.*[5] it is possible to perform relativistic XWR, and thus, relativistic, as well as electron correlation effects, can be described.
Here we present the results of HAR carried out for an organo-gold(I) compound against Mo data collected at 100 K, with the resolution limited to 0.6 Å. The outcome of DFT-based refinements with the nonrelativistic (NR) and quasi-relativistic (IOTC) approach and a few combinations of double-zeta basis sets and will be compared.

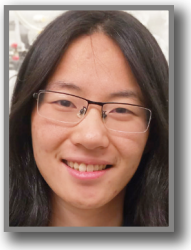
References:
1. Pyykko, P. Chem. Rev. 1988, 88 (3), 563–594.
2. Yatsimirskii, K. B. Theor. Exp. Chem. 1995, 31 (3), 153–168.
3. Hansen, N. K.; Coppens, P. Acta Crystallogr. Sect. A 1978, 34 (6), 909–921.
4. Woińska, M.; Jayatilaka, D.; Dittrich, B.; Flaig, R.; Luger, P.; Woźniak, K.; Dominiak, P. M.; Grabowsky, S. ChemPhysChem 2017, 18 (23), 3334–3351.
5. Bučinský, L.; Jayatilaka, D.; Grabowsky, S. J. Phys. Chem. A 2016, 120 (33), 6650–6669.

Overcoming distrust in solid state simulations: the case of cell parameters

Francesca Peccati[1], R. Laplaza[1,2], J. Contreras-García[1]
[1]Sorbonne Univ., CNRS, Laboratoire de Chimie Théorique, Paris, France
[2]Departamento de Química Física, Univ. de Zaragoza, Zaragoza, Spain

POSTER 31 QC
"rising star"

X-ray and neutron diffraction are well-established techniques for structural determination, whose success allowed the development of materials science. These methods are able, among other things, to provide fundamental quantities such as the values of cell parameters of crystals with an uncertainty as small as some parts in 10^4-10^5. At the same time, simulation techniques are providing with each passing day a deeper insight into the structure and properties of materials. Ideally, from now on simulation and experiment will walk hand in hand, sharing goals and fields of study. An obstacle to this cooperation is the common lack of a degree of uncertainty associated with data that are the product of computation. Without an appropriate error bar, direct comparison of the calculated value with the corresponding experimental quantity can lack physical significance. In this contribution we focus on the most basic property of a solid, the geometry of its unit cell, and employ the known delocalization error of DFT and HF methods to develop a simple and robust procedure to quickly estimate the error bar associated to calculated cell parameters. We suggest that this method should be added to the standard toolkit of solid state simulations as its validity has been proven to extend over several classes of crystals (ionic, covalent, molecular).



POSTER 32 QC

Metal-Organic Framework (MOF) as a platform for crystallographic study on small molecules

Xiaokun Pei[1,2,3,4], E. Kapustin[1,2,3,4], O.M. Yaghi[1,2,3,4,5]
[1]Dept of Chemistry, Univ. of California–Berkeley, Berkeley, CA, USA
[2]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3]Kavli Energy NanoSciences Inst. at Berkeley, Berkeley, CA, USA
[4]Berkeley Global Science Inst., Berkeley, CA, USA
[5]Center for Research Excellence in Nanotechnology, King Fahd Univ. of Petroleum and Minerals, Dhahran, Saudi Arabia

Single-crystal X-ray diffraction analysis is a powerful technique to determine the spatial arrangement of molecules within crystal structures, but in many cases crystallization is hard to achieve for flexible structures, reactive species, and impure mixtures. To overcome this challenge, in 2016, a new approach termed coordinative alignment method was reported, where the guest molecules are coordinatively attached to the backbone of a chiral metal-organic framework (MOF) to align them into an ordered pattern suitable for crystal structure determination. The structure of the host framework, MOF-520, is constructed from $Al_8(OH)_8(HCOO)_4(\text{-}COO)_4$ secondary building units (SBUs). Within these SBUs four bridging formate groups can be replaced with molecules containing either carboxylates, primary alcohols, phenols and vicinal diols. It has been demonstrated to be a potentially useful tool for solving structures of unknown molecules: sixteen molecules containing these four functional groups, ranging in complexity from methanol to plant hormones (gibberellins, containing eight stereocenters), were successfully crystallized and had their precise structure determined. However, the method limited to only a few specific functional groups. To address this problem, we aim to expand the scope of functional groups that can be used in the coordinative alignment method. Specifically, the following groups functional groups were investigated in this project: (i) pyrazolate and its analogues, (ii) sulfate/sulfonate/sulfinate, (iii) phosphate/phosphonate/phosphinate. All of these functional groups can be anchored to the Al-based SBU of MOF-520, and the structure-unknown molecules with these functionalities are incorporated and their structures are studied. Detailed insight into the selectivity of chiral incorporation encourage us to analyze the asymmetric orbital arrangement in the chiral extended framework.

Electronic properties and bonding in azolate based coordination polymers

Stefano Racioppi[1], V. Colombo[1], A. Sironi[1], M. Andrzejewski[2], P. Macchi[2]
[1]Univ. degli Studi di Milano, Dip. di Chimica, Milan, Italy
[2]Univ.of Bern, Dept of Chemistry and Biochemistry, Bern, Switzerland

POSTER 33 QC
"rising star"

Materials having dielectric constant lower then silicon dioxide (≈ 4) as interlayer insulators in modern microelectronic devices are nowadays necessary to guarantee high performance and low power consumption. In order to reach lower and lower dielectric constants the research moved to porous materials, in particular to Metal Organic-Frameworks (MOFs), which are probably the most promising systems as next-generation insulators. However, mechanical and thermal stability up to 400 °C is demanded for the application in electronic devices. In this regard, Metal Azolate Frameworks (MAFs) are chemically and thermally more stable then classic systems based on carboxylic ligands and are indeed more suitable to overcome stability challenges. In order to verify and quantify the performances of azolate based MOFs for their use as low dielectric constant materials, we investigated the Cu_4L_4·2EtOH (L=5-(4-pyridyl) tetrazolate) system. Susceptibility and dielectric constant of the evacuated MOF were obtained using Crystal14 code, polarizability with and without solvent interactions were computed via DFT calculation in gas-phase. Theoretical analysis were supported by high resolution x-ray diffraction experiment and using Hansen&Coppens multipolar model.
Moreover, the decoration of L with one fluorine atom (LF=5-(2-fluoro-4-pyridyl) tetrazolate) produced two new isostructural coordination polymer with coinage metal cations: AgLF and CuLF. Chemical behavior under pressure of both porous (Cu_4L_4·2EtOH) and not-porous (AgLF, CuLF) coordination polymers was studied. In the first case interlayer connections were induced as consequence of the extension of coordination character of the ligand, meanwhile in the second case the formation of an argentophilic interaction was observed, but not a cuprophilic one. Topological analysis were carried out with TOPOND package implemented in Crysta14, energy decomposition analysis were computed in gas-phase with EDA and IQA methods.

Can be the ligand selectivity phenomenon explained with electron density approach?

Katarzyna Rzęsikowska, J. Kalinowska-Tłuścik[1], A. Krawczuk[1]
Dept of Crystal Chemistry and Crystal Physics, Faculty of Chemistry, Jagiellonian Univ, Kraków, Poland

POSTER 34 QC

Predicting the selectivity profile of molecules acting on monoamine G-protein coupled receptors (GPCRs) is a critical issue in psychiatric drug design. However, the task is very complicated and still little is known about pharmacodynamic properties of psychoactive molecules which could allow selective binding to chosen target. 5-HT_{1A} and 5-HT_7 receptors are one of the most studied subtypes of serotonin receptors in the central nervous system. They play an important role in functioning of human body (e.g. circadian rhythm, cognition or emotions). On the molecular basis, activation of these GPCRs release an opposite effect on adenylyl cyclase function. Despite knowledge of the putative binding modes and large number of known ligands of 5-HT_{1A}R and 5-HT_7R, drugs currently used in therapies related to serotoninergic system dysfunction are, in most cases, not selective. They usually influence various monoamine targets, affecting level of other neurotransmitters, e.g. dopamine. In our research, an attempt to correlate the ligand's electron density distribution with predicted active geometry of small molecules inside the receptor's binding site was performed. This investigation has been undertaken in order to find the explanation of selectivity phenomenon of the studied ligand-receptor systems. For this purpose compounds with confirmed affinity to 5-HT_{1A} and 5-HT_7 receptors were selected form ChEMBL database (ver. 22.1). Homology models of 5-HT_{1A} and 5-HT_7 receptors were constructed with MODELLER (9.14) software, using the X-ray structure of β2-adrenergic receptor (PDB ID:2RH1) as a template. The obtained models has been evaluated using virtual screening approach. Docking study (GOLD ver. 5.4.0) for selected ligands was performed, followed by calculation of topological and energetic properties of the electron density distribution at geometries from docking poses. This approach was applied in order to define differences in the binding modes and selective ligand properties.

Experimental charge density studies of short strong hydrogen bonds (SSHBs) with potential proton migration behaviour on I19, Diamond Light Source

Lucy K. Saunders[1], C.S. Frampton[2], H. Nowell[1], D.R. Allan[1]
[1]Diamond Light Source, Harwell Science and Innovation Campus, Didcot, UK
[2]Wolfson Centre for Materials Chemistry, Brunel Univ. London, Uxbridge, UK

POSTER 35 QC
"rising star"

Short strong hydrogen bonds (SSHBs) are those with very short donor-acceptor distances in the range of 2.5 Å for O—H···O HBs[1] and 2.6 Å for N—H···O HBs.[2] The short donor-acceptor distances perturbate properties of molecules in the solid state[3] resulting in interesting physical or chemical behaviour including ferroelectrics[4] or thermochromism.[5,6] This behaviour occurs where the short donor-acceptor distance results in a potential energy surface where the barrier to proton transfer is low.[7] In this situation, the hydrogen atom may transfer across the SSHB, typically from the donor to the acceptor or into a more central position.[8,9]

In this work, we use experimental charge density studies to characterise SSHBs in potential temperature dependent proton migration materials with the aim of increasing our understanding of where such behaviour may be manifest. We present results from multipolar refinements (performed in XD2006[10] using Hansen and Coppens formulism[11]) of charge density data collected on a number of systems on beamline I19, Diamond Light Source (UK). We also discuss aspects of the data collection strategy implemented to obtain the quality high resolution data required for charge density analysis.

1. P. Gilli, et al., J. Am. Chem. Soc., 1994, 116, 909-915.
2. T. Steiner, I. Majerz and C. C. Wilson, Angew. Chem. Int. Ed., 2001, 40, 2651-2654.
3. F. Fontaine-Vive et al., J. Am. Chem. Soc., 2006, 128, 2963-2969.
4. S. Horiuchi and Y. Tokura, Nat. Mater., 2008, 7, 357.
5. D. M. S. Martins et al., J. Am. Chem. Soc., 2009, 131, 3884-3893.
6. C. L. Jones, C. C. Wilson and L. H. Thomas, CrystEngComm, 2014, 16, 5849-5858.
7. C. L. Perrin, Acc. Chem. Res., 2010, 43, 1550-1557.
8. A. O. F. Jones et al., Cryst. Growth Des., 2013, 13, 497-509.
9. J. A. Cowan, J. A. K. Howard, G. J. McIntyre, S. M. F. Lo and I. D. Williams, Acta Crystallogr. Sect. B, 2005, 61, 724-730.
10. A. Volkov et al., XD2006, 2006.
11. N. K. Hansen and P. Coppens, Acta Crystallogr. Sect. A, 1978, 34, 909-921.

POSTER 36 QC
"rising star"

Electron density and dielectric properties of highly porous MOFs

Rebecca Scatena, Y.T. Guntern, P. Macchi
Dept of Chemistry and Biochemistry, Univ. of Bern, Bern, Switzerland

Porous metal organic frameworks (MOFs) have been addressed as promising next generation low dielectric constant (ε_r) materials[1]. The strategy to reach values lower than ε_r 2.5 consists in the partial replacement of dense material with the lowest possible ε_r contribution, which is air or vacuum[2]. The investigation of structure/property relationship allows rational selection and design of operating materials among the virtually infinite plethora of MOFs[3,4]. Our approach consists in the correlation between electron density distribution of isostructural MOFs and their dielectric properties, both computed and measured. The selected materials are $[Cu_3(BTC)_2]$, also known as Hkust-1, and $[Zn_3(BTC)_2]$ which present a 3D system of pores accounting for 69 % of the volume[5,6]. We have determined the electron density distribution from single crystals with various degrees of pores activation. Furthermore, quantum mechanical simulations have been used to rationalize the networks polarization induced by guest molecules. The derived bonding properties have been correlated to atomic and functional group polarizabilities. By measuring the dielectric constant over a wide range of frequencies for differently activated samples, it has been possible to show the role of pores morphology and content. This study has highlighted the main factors which tune dielectric properties in highly porous materials and reports, to the best of our knowledge, the first comparison between calculated and experimental ε_r of MOFs.

1. Usman, M., *et al.*, *ChemElectroChem* **2,** 786–788 (2015)
2. Zagorodniy, K., *et al.*, *Appl. Phys. Lett.* **97,** 2013–2015 (2010)
3. Yaghi, O. M., *et al.*, *Nature* **423,** 705–714 (2003)
4. Wang, Z., *et al.*, *Chem. Soc. Rev.* **38**, 1315-1329 (2009)
5. Wu, Y., *et al.*, *Angew. Chemie - Int. Ed.* **47,** 8929–8932 (2008)
6. Feldblyum, J. I., *et al.*, *J. Am. Chem. Soc.* **133,** 18257–18263 (2011)

[Withdrawn]

POSTER 37 QC

POSTER 38 QC

Structural phase transition in iron borate $HoFe_3(BO_3)_4$

Ekaterina Smirnova[1], O. Alekseeva[1], A. Dudka[1], D. Khmelenin[1], I. Gudim[2], L. Bezmaternykh[2], K. Frolov[1], I. Luybutin[1]
[1]Shubnikov Inst. of Crystallography of FSRC "Crystallography and Photonics" RAS, Moscow, Russia
[2]Inst. of Physics of Siberian Branch, Russian Academy of Sciences, Krasnoyarsk, Russia

A number of phase transitions was discovered in $HoFe_3(BO_3)_4$ compound which belong to a multiferroic family $RFe_3(BO_3)_4$. Magnetic ordering is observed below 38 K and spin-reorientation take place at about 5 K [1]. There is a structural phase transition sp. gr. $R32 \rightarrow P3_121$ in $HoFe_3(BO_3)_4$ at 427 K according to DTA data for powder samples [2] and at 360 K according to high-resolution spectroscopy for single crystals grown with $Bi_2Mo_3O_{12}$ in a flux [3].
In the present work $HoFe_3(BO_3)_4$ single crystals grown with $Bi_2Mo_3O_{12}$ in a flux were studied by X-ray diffraction analysis in a temperature range 11–500 K.
X-ray diffraction measurements were performed using synchrotron radiation at PILATUS@SNBL diffractometer and using a laboratory diffractometer Xcalibur CCD (Oxford Diffraction). The chemical composition of the compound was verified by X-ray energy-dispersive elemental analysis.
Bi atoms from the flux were discovered in the structure. According to X-ray diffraction and energy-dispersive elemental analysis, Bi atoms partly occupy about 5% of Ho atomic position. The temperature dependence of unit cell parameters *a,b* demonstrates a sharp decrease at about 365-370 K with temperature lowering, which responds to the structural phase transition. An analysis of reflections forbidden in high-temperature sp. gr. *R*32 showed that the number and intensity of such reflections increase sharply below 365 K, which allowed us to consider T_s=365 K. Crystal structure of $HoFe_3(BO_3)_4$ was refined for several data sets at temperatures 90–500 K. With temperature lowering a non-uniform change in the bond lengths is observed in the $Ho(Bi)O_6$ prisms, $B2O_3$ and $B3O_3$ triangles, and FeO_6 octahedra.

The study was partly supported by the RFBR (pr.no. 17-02-00766 A).

1. C. Ritter et al. / J. Phys.: Condens. Matter 20 (2008) 365209 (9pp)
2. Y. Hinatsu et al. / Journal of Solid State Chemistry 172 (2003) 438–445
3. D. A. Erofeev et al. / Optics and Spectroscopy 120, No. 4 (2016), pp. 558–565

Electron densities of organic molecular crystals from powder X-ray diffraction

Bjarke Svane, B. Brummerstedt Iversen
Dept Chemistry, Aarhus Univ., Aarhus, Denmark

POSTER 39 QC
"rising star"

Detailed knowledge of the nature of the chemical bonding is a prerequisite for understanding the physical and chemical properties of materials and this information is best available in the electron density (ED). Virtually all experimental ED distributions are determined from structure factors extracted from single crystal X-ray diffraction, since this has been regarded the optimal way to obtain data of the highest quality. However, our recent work has shown that data obtained from powder X-ray diffraction (PXRD) can exceed the data quality from single crystal diffraction[1-4]. At the same time PXRD is experimentally less demanding and time-consuming[5].

I aim to determine ED distributions and atomic displacement parameters (ADPs) of molecular materials from highly accurate PXRD data using Hirshfeld Atom Refinement (HAR) and multipole modelling. HAR is similar to other refinement methods with the crucial difference that the atomic scattering factors are obtained from aspherical atom partitioning from an ab initio quantum mechanical model[6]. In this way, it is possible to obtain more precise molecular geometries and ADPs than possible from conventional refinement methods[7]. The better model will lead to more accurate intensity extraction, which can also be used in a subsequent full ED refinement using the multipole method. This is of critical importance if the PXRD method to obtain EDs is to be extended from small unit cell inorganic solids to molecular crystals with severe peak overlap.

[1] N. Bindzus et al. (2014), Acta Cryst. A70, 39-48
[2] T. Straasø et al. (2014), J. Synchrotron Rad., 21, 119-126
[3] N. Wahlberg et al. (2015), J. Phys. Chem. C, 119, 6164-6173
[4] K.Tolborg et al. (2017), Acta Cryst. B73, 521-530.
[5] M. R. V. Jørgensen et al. (2014), IUCrJ, 1, 267-280
[6] S. C. Capelli et al. (2014), IUCrJ, 1, 361-379
[7] M. Fugel et al. (2018), IUCRJ, 5, 32-44

POSTER 40 QC
"rising star"

Application of HAR to incomplete, high pressure data

Daniel Tchoń, A. Makal
Faculty of Chemistry, Univ. of Warsaw, Warsaw, Poland

Despite the rapid development of hardware, software and methods utilised in the field of quantum crystallography, high-pressure (HP) diffraction experiments remain a very challenging subject in terms of precise electron density reconstruction. Diamond Anvil Cells (DACs) used to pressurise the sample heavily limit an accessible volume of reciprocal space, causing the reflection data to be systematically incomplete. A Fourier transform performed over a limited "disc" of reflection data results in a charge topology being heavily distorted, rendering multiple standard modelling methods inapplicable.

Hirschfeld Atom Refinement (HAR) is a relatively new refinement technique implemented in Tonto software, which allows one to precisely retrieve as much structural information as possible from a single X-ray diffraction experiment. Recent research by Woińska et al. show that HAR is capable of precise and accurate hydrogen atom position and displacement characterisation even for data trimmed to a low resolution. Based on this fact, preliminary calculations on artificially de-completed datasets for α-glycine and 1-phenyl-5-(1'-pyrene)-pyrazole (both crystallising in monoclinic system, space group $P2_1/n$) have been performed. Although HAR requires supposedly very redundant and complete data, the refinements against heavily trimmed reflection sets, trunctuated both randomly and systematically, have been successful. While a meticulous analysis of output models is still due, they feature reasonable hydrogen atom behaviour and are outwardly indistinguishable from respective references.

Hirschfeld Atom Refinement is unlikely to consistently yield models suitable for reliable topological analysis based on HP data collected on low-symmetry systems. It does, however, appear to be a promising alternative to the constrains-supported Independent Atom Model.

This study was financially supported by the Polish National Science Centre (NCN) based on decision UMO-2015/17/B/ST4/04216.

Charge density analysis of Lamivudine nitrate (3TC), an anti-HIV drug

Juan Tenorio[1,2], J. Ellena[1], C. Lehmann[2]
[1]São Carlos Inst. of Physics, Univ. of São Paulo, São Carlos, Brazil
[2]Max-Planck-Inst. für Kohlenforschung, Mülheim a.d. Ruhr, Germany

POSTER 41 QC

Theoretical and experimental studies for the determination of the molecular Electron Density (ED) of crystalline organic compounds, mainly of compounds with pharmaceutical activity, have recently drawn the attention in the area of small-molecule crystallography and pharmaceutical science.[1] In that vein, here we report a crystal engineering study and Charge Density Analysis (CDA) of a salt of Lamivudine (3TC), one of the well marketed and successful Nucleoside Reverse Transcriptase Inhibitors (NRTI). This study aimed to assess, from the viewpoint of the molecular electron density distribution, the physicochemical stability and the supramolecular properties of the nitric acid salt derivative (anhydride Lamivudine nitrate, 3TCH-NO_3) since the 3TC active pharmaceutical ingredient used in the market has stability problems in the manufacturing process.[2,3]
Experimental CDA was determined by high resolution X-ray diffraction data ($\sin\theta_{max}/\lambda_{Mo} \approx 1.2$ Å^{-1}) through multipolar refinement based on the Hansen-Coppens formalism, and it is compared with a model obtained by refinement of theoretical structure factors calculated by periodic theoretical calculations at the B3LYP/6-311G* level of theory. Features of the molecular conformation, topological analysis by quantum theory of atoms in molecules (QTAIM), supramolecular behavior and physicochemical stability of this new solid form are assessed and correlated with the CD distribution. Therefore, we aim to shed light on the pharmaceutical properties of a new solid form of the 3TC drug by a combination of crystal engineering and CD studies, which is a new approach in the analysis of a pharmaceutical compound.

1. Krawczuk, A. & Macchi, P. (2014). Chemistry Central Journal. 8:68.
2. Martins, F. T. et al. (2012). J. Pharm. Sci. 101 (6), 2143–2154.
3. Perumalla, S. R. & Sun, C. C. (2014). Cryst. Growth Des. 14 (8), 3990-3995.

POSTER 42 QC

Charge densities from powder X-ray diffraction

Kasper Tolborg[1], M.R.V. Jørgensen[1,2], H. Kasai[1,3], J. Becker[1], A.-C. Dippel[4], J. Als-Nielsen[5], B.B. Iversen[1]
[1]Center for Materials Crystallography, Dept of Chemistry and iNANO, Aarhus Univ., Aarhus, Denmark
[2]MAX IV Laboratory, Lund Univ., Lund, Sweden
[3]Faculty of Pure and Applied Sciences, CiRfSE and TIMS, Univ. of Tsukuba, Tsukuba, Ibaraki, Japan
[4]PETRA III, Deutsches Elektronen-Synchrotron, Hamburg, Germany
[5]Niels Bohr Inst., Univ. of Copenhagen, Copenhagen, Denmark

Traditionally Single Crystal X-Ray Diffraction (SCXRD) is used to obtain structure factor amplitudes for Electron Density (ED) modelling in crystalline materials. However, for SCXRD there are important issues with extinction, absorption and scaling between detector frames. All three issues can be partially solved by performing Powder X-Ray Diffraction (PXRD). Two major issues are present in PXRD; peak overlap and background treatment.
Recently, we have shown that PXRD can be a valuable alternative to SCXRD for ED determination in high symmetry inorganic materials, and subtle deformations in core ED have been determined for diamond, silicon and cubic boron nitride. For silicon and diamond, the extracted structure factors were shown to be of comparable precision and accuracy to the dynamical *Pendellösung* data [1,2]. This was made possible by performing diffraction in vacuum with the first version of our Aarhus Vacuum Imaging plate Diffractometer with a sample-to-detector distance of 300 mm. We have now finished commissioning of a new version of AVID with sample-to-detector distance increased to 1200 mm, which is shown to give a large improvement in signal-to-background ratio and peak broadening due to the differences in radial dependences of coherent and incoherent scattering [3]. Benchmark PXRD data on silicon has been collected on the new AVID with resolution up to $\sin\theta/\lambda > 2.0$ Å^{-1}. Data are modelled using a combined HC-Rietveld method, where overlapping reflections are partitioned based on the Hansen-Coppens model, which reduces bias towards the independent atom model.
The extracted structure factors are shown to be of improved precision and accuracy at high angles. These are used for modelling the ED in silicon including deformations in the core ED with improved accuracy compared to previous studies.

[1] Bindzus, N. et al. (2014). Acta Cryst. A70, 39-48
[2] Wahlberg, N. et al. (2016). Acta Cryst. A72, 28-35
[3] Tolborg, K. et al. (2017). Acta Cryst. B73, 521-530

POSTER 43 QC

QTAIM theoretical studies on manganese(I), ruthenium(I), rhodium(I), and iridium(I) borane complexes.

Juan Francisco Van der Maelen[1], J.A. Cabeza[2]
[1]Dept Physical and Analytical Chemistry, Univ. Oviedo, Asturias, Spain
[2]Dept Inorganic and Organic Chemistry, Univ. Oviedo, Asturias, Spain

In their transition-metal complexes, amine–boranes and aminoboranes use one or two of their BH groups to bind the metal atom (σ-complexes), implicating only the H atom (Shimoi-type coordination) or both B and H atoms (agostic-type coordination) of the B–H bond. Theoretical QTAIM gas-phase studies have shown that the attachment of the BH_3 group to the metal atom in the octahedral complex [Mn(κ^3N,H,H-*mapy*BH_3)$(CO)_3$] (H*mapy* = 2-(methylamino)pyridine) is symmetric, and it involves two B–H–M interactions that are intermediate between Shimoi and agostic types.[1] The symmetric structure found for the above complex also supports the proposal that the asymmetric coordination found for the BH_2 fragment in the Ru complex [RuH(κ^3N,H,H-*mapy*BH_3)(CO)(P^iPr_3)] is not due to its octahedral coordination geometry, which is present in both Mn and Ru complexes, but to the different trans-influence of the hydride and phosphane ligands (which are trans to H atoms in the BH_2 group). In the Ru complex, the attachment of both borane B–H bonds to the metal atom is also intermediate between those of Shimoi and agostic types, but the B–H bond trans to the hydride interacts more agostically with the metal atom than the B–H bond that is trans to the phosphane. Moreover, these results are complementary to those found on the trigonal bipyramidal complexes [*M*(κ^3N,H,H-*mapy*BH_3)(*cod*)] (*M* = Rh, Ir) (cod = cycloocta-1,5-diene) where the coordination of the BH_2 fragment is also asymmetric in both, but not because it is influenced by the other ligands in these instances but because it occupies two equatorial coordination sites whose ideal coordination angle (120°) is too wide to efficiently accommodate the BH_2 fragment in a symmetric manner.[2]

[1]J. Brugos, J. A. Cabeza, P. García-Álvarez, E. Pérez-Carreño, and J. F. Van der Maelen, Dalton. Trans., 2017, 46, 4009-4017.
[2]J. Brugos, J. A. Cabeza, P. García-Álvarez, A. R. Kennedy, E. Pérez-Carreño, and J. F. Van der Maelen, Inorg. Chem., 2016, 55, 8905-8912.

POSTER 44 QC
"rising star"

Electronegativity in band structure calculations

Alena Vishina, Martin Rahm
Chalmers Univ. of Technology, Gothenburg, Sweden

Band structures in electron structure calculations are often too cumbersome to analyse. We are trying to simplify this analysis by looking for some easy-to-calculate descriptors that can be obtained from the band structures, such as electronegativity. We introduce and investigate the concept of elecronegativity resolved in reciprocal space by calculating it from the band structure for various materials such as metals, semiconductors, ionic solids, etc.

On the compressibility of C-H bonds in late transition metal alkyls

Marcel Vöst, D. Schmitz, M. Kalter, W. Scherer
Dept Physics, Augsburg Univ., Augsburg, Germany

POSTER 45 QC
"rising star"

For investigating the pressure dependency of C-H bonds in organometallic transition metal alkyls, the chlorotrimethylplatinum(IV)-complex $[PtCl(CH_3)_3]_4$ (**1**) has been selected as a reference system to study the compressibility of organometallic *M*-C and C-H bonds while C-H bond activations by the proximity of the metal center are hindered. The cubic symmetry of **1** provides a virtually isotropic compression scenario and well-resolved, non-convoluted Raman spectra can be obtained over the studied pressure range of 0 to 4.9(1) GPa. Hence, the shift of the symmetric C-H stretching mode reveals pressure-dependent changes of the C-H bond strength and can be correlated with the according bond length. All normal modes of the Raman spectra could be assigned by employing theoretical frequencies based on DFT calculations of **1**. At 4.9(1) GPa a significant blue shift of the symmetric C-H and Pt-C stretching mode can be observed in the Raman spectrum, indicating a shortening of the respective bond length. To monitor the unit cell compression and structural changes of **1**, X-ray diffraction studies have been carried out in the pressure range of 0 to 4.9(1) GPa. The determined low bulk modulus k_0 = 7.4(6) GPa of **1** results from a high intermolecular compressibility and strongly decreasing intramolecular metal-to-ligand distances (Pt-Cl). For the Pt-Cl bond the X-ray diffraction data reveal a significant reduction of the bond length up to 4.9(1) GPa. In contrast, the pressure dependency of the Pt-C bond suggests only a trend of a reducing bond length, wherein the changes stay within the estimated standard deviation. However, this trend is supported by the observed significant blue shift of the Pt-C stretching mode and symmetric C-H mode in the Raman spectrum, ruling out any C-H bond activation by agostic interactions.

POSTER 46 QC

HAR and TAAM refinements against CuKα and MoKα X-ray diffraction data

Monika Wanat[1,2], M. Malińska[2], K. Woźniak[2]
[1]College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences (MISMaP), Univ. of Warsaw, Warsaw, Poland
[2]Biological and Chemical Research Centre, Chemistry Dept, Univ. of Warsaw, Warsaw, Poland

Commonly, in the case of routine X-ray data refinement, the Independent Atom Model (IAM) of electron density is used. Nonetheless, atoms are assumed to be neutral and spherical, a quantitative description of electron density distribution is not given. A far better model that allows for modelling of deformation of spherical charge density was introduced by Hansen and Coppens[1] and is called a pseudoatom model of electron density. Application of this model requires an excellent quality crystals and high resolution XRD data. Quite often, this is difficult to be fulfilled. Therefore, new methods have been developed that enable reconstruction of electron density i.e. Hirshfeld Atom Refinement (HAR)[2] or Transferable Aspherical Atom Model (TAAM)[3].
We will show that CuKα X-ray Diffraction Data could be refined with HAR or TAAM methods. HAR and TAAM refinements for model crystal structures using CuKα and MoKα X-ray Diffraction Data will be compared. Hydrogen atoms in these refinements were treated in various ways: isotropic, anisotropic or estimated using Shade. Additionally, for MoKα X-ray diffraction data, the multipole model and high order TAAM and HAR refinements will be presented. Obtained results will be shown in comparison with neutron diffraction data. Particularly, the ADPs analysis using Peanut show that ADPs of hydrogen atoms are better modelled using CuKα X-ray data. Moreover, analysis of geometry, fractal dimension plots and residual density maps will be shown.

References:
(1) Hansen, N. K.; Coppens, P. Acta Crystallogr. A 1978, 34 (6), 909–921.
(2) Capelli, S. C.; Bürgi, H.-B.; Dittrich, B.; Grabowsky, S.; Jayatilaka, D. IUCrJ 2014, 1 (5), 361–379.
(3) Jarzembska, K. N.; Dominiak, P. M. Acta Crystallogr. A 2012, 68 (1), 139–147.

Heavy meets light - A systematic study of the bond between hydrogen and heavy elements

Erna K. Wieduwilt[1], L.A. Malaspina[1], D. Duvinage[1], M. Olaru[1], A.J. Edwards[2], J. Beckmann[1], D. Jayatilaka[3], S. Grabowsky[1]
[1]Univ. of Bremen, Inst. of Inorganic Chemistry and Crystallography, Bremen, Germany
[2]Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW, Australia
[3]Univ. of Western Australia, School of Molecular Sciences, Perth, WA, Australia

POSTER 47 QC
"rising star"

The determination of hydrogen atom parameters from X-ray diffraction experiments is still a challenge because its X-ray scattering power is low and its single electron is shifted towards the bond. This makes a spherical atom approach inadequate for its description. Bonded to heavy atoms, hydrogen atoms are even more difficult to localize, as Fourier truncation ripples close to the heavy atom positions can hide the weak electron density maxima representing the hydrogen atoms.

There are several ways how hydrogen atom positions can be treated in general within a structure refinement[1], out of which two are closer examined in this study: the Hirshfeld Atom Refinement[2] (HAR) and neutron diffraction. HAR can be performed using routine in-house X-ray data. A statistical validation of HAR has previously been performed on small organic molecules and shows good agreement between neutron and HAR element-hydrogen bond lengths[3]. An expansion to inorganic molecules has not been possible, since even for main-group element E-H bonds there is an astonishing lack of accurate low-temperature neutron-diffraction data available in the literature[3].

This study has two main objectives: (i) acquiring neutron derived geometries to be used as reference values to validate the (ii) HAR for compounds involving H atoms next to heavy elements. We will present the results of the neutron diffraction experiments collected at the Bragg Institute of the Australian Nuclear Science and Technology Organisation (ANSTO) at the instrument KOALA using the Laue technique on compounds involving E-H bonds, as well as some preliminary results and challenges encountered during the HAR on the same set of compounds.

[1]L. A. Malaspina et al. Cryst. Growth Des. 2017, 17, 3812-3825.
[2]D. Jayatilaka, B. Dittrich, Acta Cryst. A 2008, 64, 383. S. C. Capelli et al. IUCrJ 2014, 1, 361-379.
[3]M. Woinska et al. Sci. Adv. 2016, 2, e1600192.

Quantitative and qualitative analysis of intermolecular interactions of multicomponent crystals containing sulfonamides

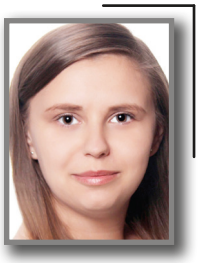
Joanna Wojnarska
Dept Chemistry, Univ. Jagiellonian, Krakow, Poland

POSTER 48 QC

Crystalline materials with desired physical properties are of major interest in modern science and technology. For rational design of new crystal phases with prospective optical properties, specific building blocks need to be used. The chosen components should have large values of (hyper)polarizability and their mutual assembly needs to maximize (non)linear optical effects. Electron density (ED) analysis allows for accurate, quantitative investigation of intermolecular interactions, which are of great importance for identification of factors that promote particular packing of building blocks. When combined with qualitative tools as Hirshfeld Surfaces (HS)[1], Non-Covalent Interaction analysis (NCI)[2], the full information about dispersive and directional interactions can be established. The obtained knowledge allows to identify reproducible synthons, which can be useful in crystal engineering of new functional materials.

The designed multicomponent materials containing sulfanilamide: sulfanilamide sulfamic acid salt and sulfanilamide [(4-sulfamoylphenyl)carbamoyl]formic acid salt have been investigated to determine factors leading to the formation of a particular crystal structure. The intermolecular interactions have been carefully studied using electron density analysis and qualitative crystal engineering methods: HS and NCI index. The topology of ED have been explored by means of Quantum Theory of Atoms in Molecules (QTAIM)[3]. Experimental ED distribution have been compared with theoretically determined values obtained from periodic calculations. The results allowed to confirm the presence of hydrogen bonds with an intermediate character between closed shell and shared shell interactions. Additionally, the influence of dispersive interactions on the crystal structure formation was established.

[1] Spackman, M. A. et al. (2009) CrystEngComm 11, 19.
[2] Contreas-Garcia, J. et al. (2011) J. Chem. Theory Comput. 7, 625.
[3] Bader, R. (1994) Oxford University Press.



POSTER 49 QC

N-oxide – *N*-oxide interactions and Cl…Cl halogen bonds in pentachloropyridine *N*-oxide: the many-body approach to interactions in the crystal state

Kinga Wzgarda-Raj[1], A.J. Rybarczyk-Pirek[1], M. Palusiak[1], S. Wojtulewski[2]
[1]Theoretical and Structural Chemistry Group, Dept of Physical Chemistry, Faculty of Chemistry, Univ. of Lodz, Lodz, Poland
[2]Inst. of Chemistry, Univ. of Bialystok, Bialystok, Poland

Recently, an increased interest in pyridine *N*-oxide, due to their use. Pyridine *N*-oxides are components of antifungal, antiviral, anti-inflammatory and bacteriostatic substances, as well as being used as drugs for cancer chemotherapy [1-3].
Here, in the continuation of our research, we present the results of the synthesis of a new crystal stabilized by halogen bonds, pentachloropyridine *N*-oxide. The title compound, crystallizes in the monoclinic group $P2_1/c$ with one molecule in a general position.
In the crystal structure, molecules are linked by C - Cl…Cl halogen bonds into infinite ribbons extending along the crystallographic [100] direction. These molecular aggregates are further stabilized by very short intermolecular *N*-oxide – *N*-oxide interactions into herringbone motifs [4].
Computations based on quantum chemistry methods [5] allowed for a more detailed description of the *N*-oxide – *N*-oxide interactions and Cl…Cl halogen bonds. For this purpose, the many-body approach to interaction energy were applied.

Acknowledgments
The authors acknowledge the financial support from National Science Centre of Poland (Grant No. 2015/19/ B/ST4/01773).
The Oxford Diffraction SuperNova Dual diffractometer was funded by the EFRD in Operational Programme Development of Eastern Poland 2007-2013 via Project no: POPW.01.03.00-20-004/11.

References

[1].W.A. Denny Curr. Med. Chem. Anti-Cancer Agents 2004 4, 395– 399.
[2].J.R. Amsden, P.O. Gubbins, S. McConnell, E. Anaissie Antimicrob. Agents Chemother. 2013 57, 3420–3423.
[3].J.C. Villalobos-Rocha, L. Sánchez-Torres, B. Nogueda-Torres, A. Segura-Cabrera, C.A. Garcia-Pèrez, V. Bocanegra-Garcia, I. Palos, A. Monge, G. Rivera Parasitol. Res. 2014 113, 2027–2035.
[4].K. Wzgarda-Raj, A. J. Rybarczyk-Pirek, S. Wojtulewski, M. Palusiak Acta Cryst. 2018, C74, 113-119.
[5]. M. J. Frisch et al. 2009 GAUSSIAN09. Gaussian Inc., Wallingford, CT, USA.

POSTER 50 QC

Designer triazole-based naphthalimide ligands as contributors to the spin crossover phenomenon: investigating π stacking effects using quantum crystallography

Ningjin Zhang[1], J.A. Kitchen[2], S.J. Coles[1]
[1]Dept of Chemistry, Univ. of Southampton, Southampton, UK
[2]Dept of Chemistry, Univ. of Massey, Auckland, New Zealand

The numerous interesting properties of spin crossover (SCO) active materials, combined with the current trend to develop molecular electronics and machines, has resulted in a dramatic increase in the exploration compounds exhibiting these properties in the last few years. Increasing solid-state interactions between metal complexes is essential for controlling the nature of the SCO event. One approach to achieve this is to use supramolecular chemistry to assemble complexes into ordered arrays through non-covalent interactions[1].Some previous work has investigated the effect that hydrogen bonding and halogen bonding has on the cooperative nature of the SCO event. It is proposed that other supramolecular interactions can also alter the nature of this cooperativity and the focus of this work is on utilising $\pi\cdots\pi$ interactions to systematically modify the effect.
1,2,4-Triazole-Naphthalimide-based ligands were identified as the target for this project because of: a) their inherent ability to induce SCO in Fe(II), b) the long range ordering achieved through π-stacking and c) the interesting photophysical properties of the 1,8-naphthalimide moiety. Moreover, metal binding sites can be easily incorporated, and the π-deficient naphthalimide is ideally suited for $\pi\cdots\pi$ stacking.
While systematically varying the nature of substituents on the ligand backbone, we will use quantum crystallography methods to develop an understanding of how subtle changes in electron withdrawing/donating substituents influence the nature of interactions and accordingly how π interactions influence spin crossover properties.
The calculation of intermolecular interaction energies has provided an array of information which provides insights that we wish to develop into detailed structure function relationships and thereby increase control over the behavious of spin crossover materials.

Reference
1.Gutlich, P.; Goodwin, H. A., Spin Crossover in Transition Metal Compounds I 2004, 233, 1-47.

Additional Participants

Hans-Beat Bürgi

Przemyslaw Dera

Birger Dittrich

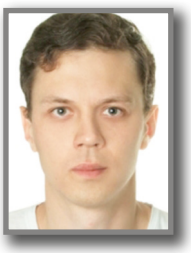
Juan Ignacio González Pérez

Regine Herbst-Irmer

Bo Iversen

Christian Jelsch

Xiaoming Jiang



Artem Kabanov

Helena Keil

Wim Klooster

Annika Münch

Presenting author index