Ultra high resolution crystallography of small molecules and proteins

Maqsood Ahmed

CRM², UMR 7036, UL, Institute Jean Barriol, Nancy, France

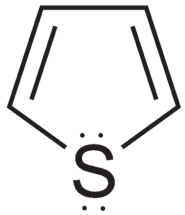
May 30th, 2012

Overview

In collaboration with

Sajida Noureen, Marc Beley, Philippe C. Gros

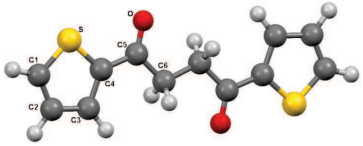
SRSMC, group Synthèse Organométallique et Réactivité (SOR), Université de Lorraine.



Thiophene

Substituted homologues of thiophenes find extensive applications in the production of:

1. Organic semiconducting materials
2. Electric field effect transistors
3. Dye-sensitized solar cells (DSSCs)*



A substituted thiophene compound

Advantages over Si based semiconductors:

1. Their electronic properties are easily tunable by chemical modifications.
2. A chemical substitution reaction is very easy to carry out.

* Grätzel, M. (2001), *Nature*. 414, 338.

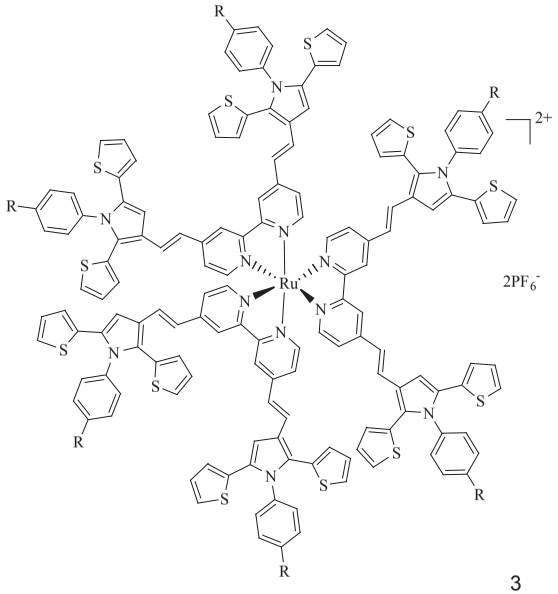
Correct knowledge about structure is crucial for a rational substitution.

- **Position of the substituents,**
- **Stereochemistry**
- **Electronic properties**

X-rays crystallography gives the most accurate structure as compared to other techniques (like NMR).

Intermolecular interactions in Thiophenes

A typical Ruthenium complex applicable in DSSCs.*



Noureen *et al.*, (2012). *Dalton Transactions*, 41, 4833.

OUTLINE

1. Structural disorder in Thiophenes
2. Charge density analysis of a Thiophene compound
3. Lone pairs electron density and H bonds
4. Transferability of electron density parameters
   - a - Small molecule
   - b - Protein
5. Conclusions & Perspectives

OUTLINE

1. Structural disorder in Thiophenes

2. Charge density analysis of a Thiophene compound

3. Lone pairs electron density and H bonds

4. Transferability of electron density parameters

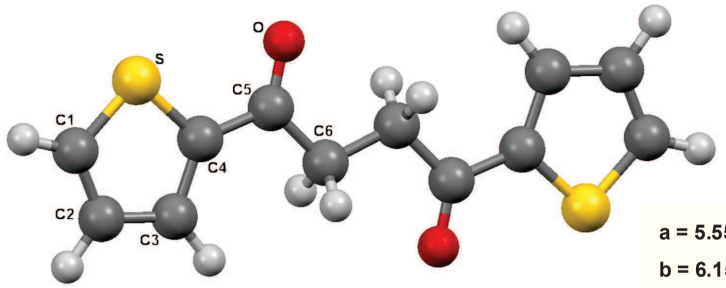
   a - Small molecule

   b - Protein

5.Conclusions & Perspectives

Disorder in *1,4-bis (2-thienyl) butane-1,4-dione*

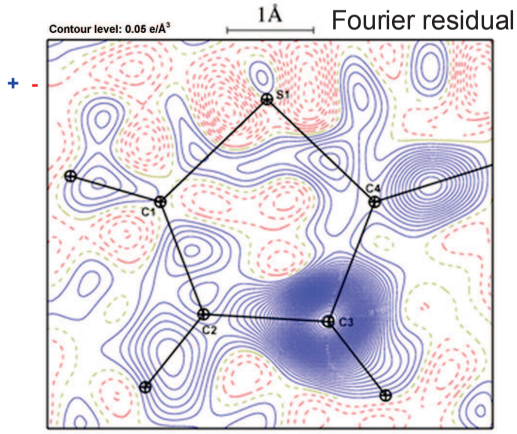
M1



Monoclinic
$P2_1/n$

a = 5.5516(5) Å β = 92.240(6)°
b = 6.1567(4) Å V = 553.3 (2) Å³
c = 16.1989(16) Å Z = 2

Temperature = 100K



Data resolution : 0.77 Å

$R(F)$ initial = 7.4 %

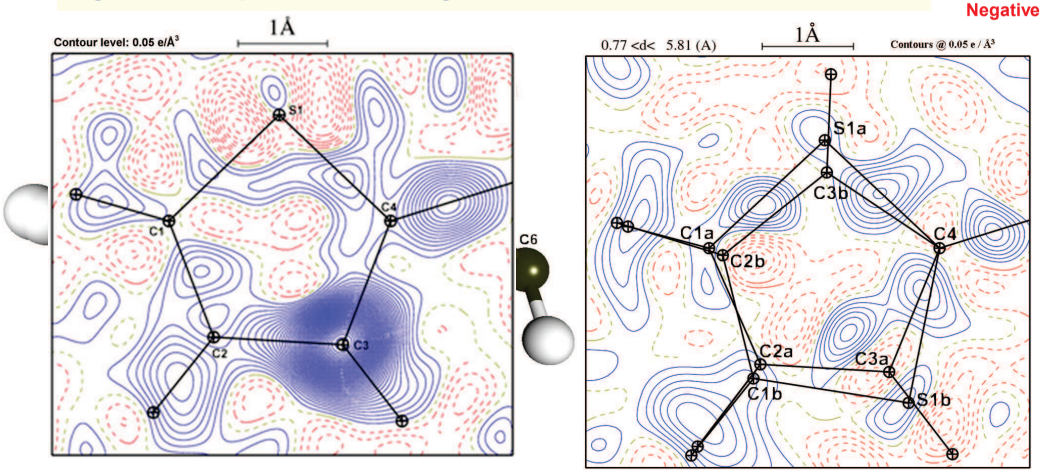
$wR^2(F)$ initial = 7.5 %

λ = 0.71073 Å

With modelling of disorder, the residual electron density shows significant improvement along with refinement statistics.

M1

Positive
Negative



$P_a + P_b = 1$

P_a = 88(3)%

P_b = 12(3)%

$R(F)$ % = 7.4 to 5.6

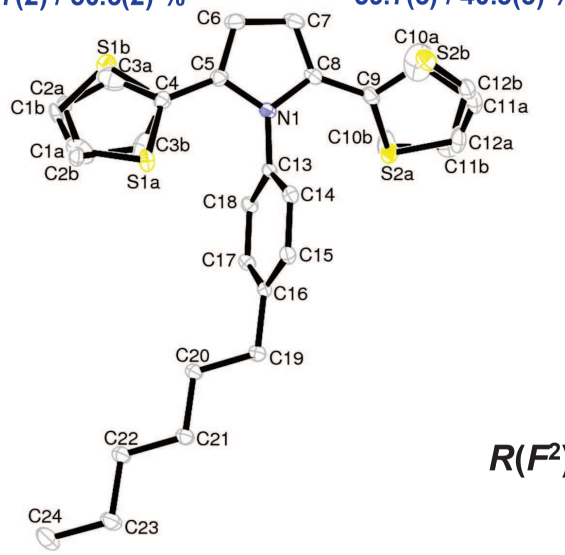
$wR^2(F)$ % = 7.5 to 4.9

Disorder in *1-(p-Hexylphenyl)-2,5-di(2-thienyl) pyrrole*

M2

49.7(2) / 50.3(2) %

59.7(3) / 40.3(3) %



Triclinic *P-1*

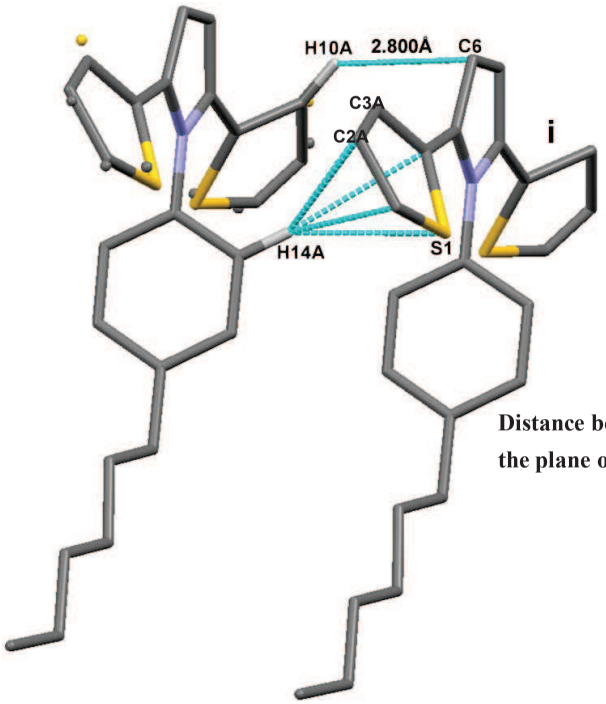
a = 5.5822(2) Å	α=64.487(4)°
b = 14.2465(5) Å	β = 92.2396(6)°
c = 14.3776(7) Å	γ= 80.637(3)
Z = 2	V = 1018.05(7)Å³

Temperature = 100K

$R(F^2)$ = 17.4 % ⟶ 5.7%

Ratio of disorder are different, even for thiophene rings within a same molecule

C-H···π interactions



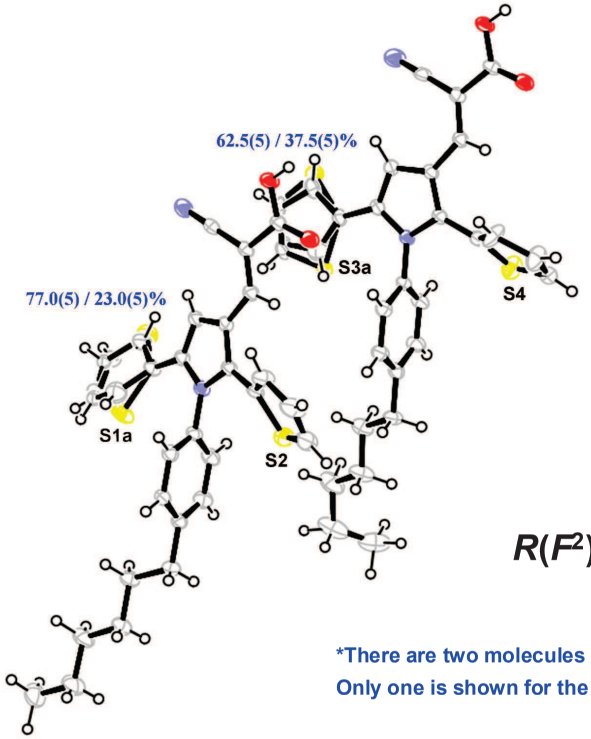
(i) = x+1, y, z

Distance between H14A atom & the plane of the ring (S1,C2A,C3A)

=

2.6Å

Disorder in 2-cyano-3-[1-(4-hexylphenyl)-2,5-di(thiophen-2-yl)-1*H*-pyrrol-3-yl]acrylic acid



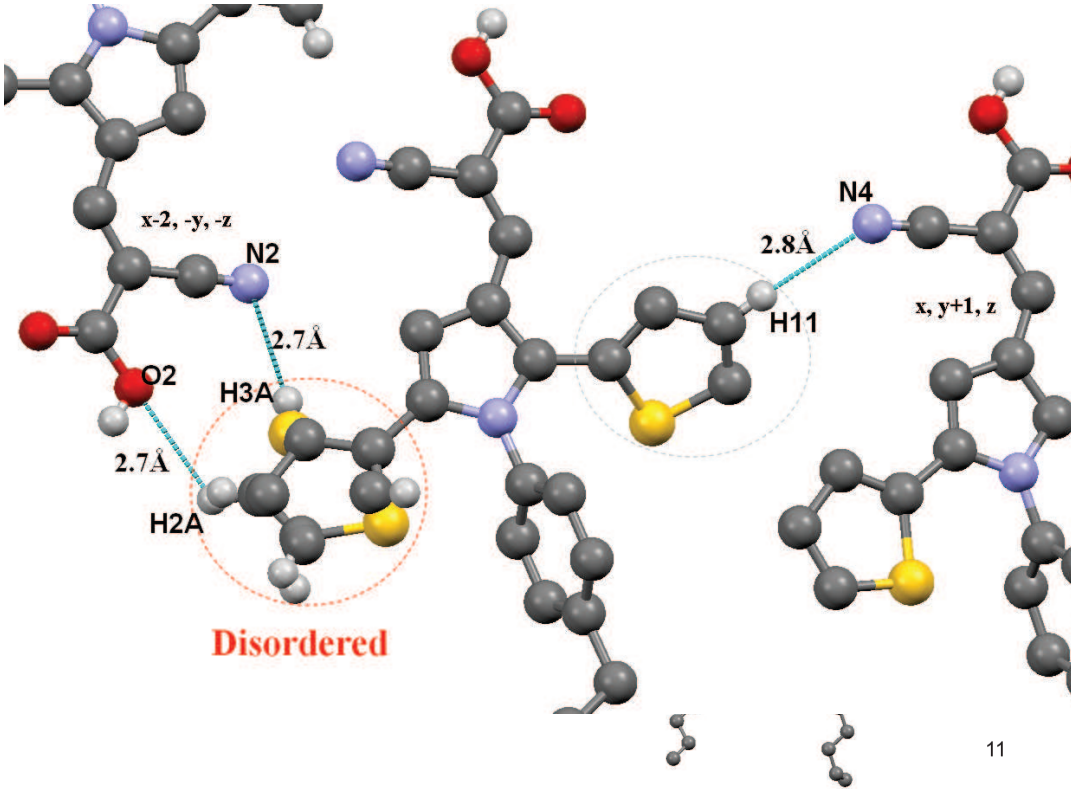
Triclinic *P-1*

a = 12.315(5) Å	*α* = 96.567(5)°
b = 12.592(5) Å	*β* = 99.016(6)°
c = 16.386(5) Å	*γ* = 96.946 (5) °
Z = 4 *Z'* = 2*	*V* = 1234.0(16) Å³

Temperature = 100K

$R(F^2)$ = 14.2 % ⟶ 6.9%

*There are two molecules in the asymmetric unit.
Only one is shown for the reason of clarity.



Conclusion

Thiophene ring shows structural disorder in the absence of a steric hindrance.

OUTLINE

1. Structural disorder in Thiophenes

2. Charge density analysis of a Thiophene compound

3. Lone pairs electron density and H bonds

4. Transferability of electron density parameters

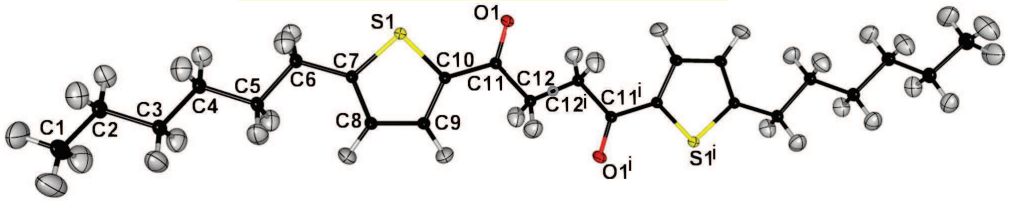
   a - Small molecule

   b - Protein

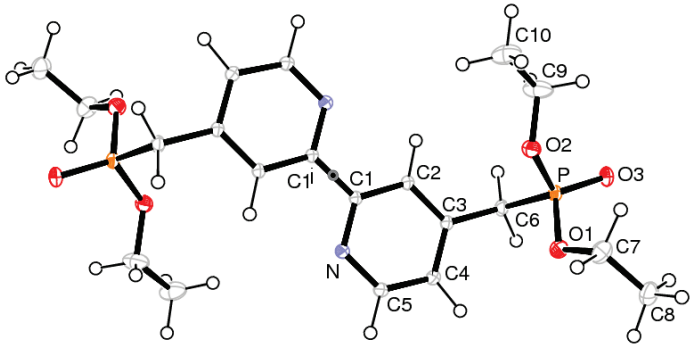
5.Conclusions & Perspectives

X-rays charge density analysis

1,4-bis (5-hexyl-2-thienyl) butane-1,4-dion



Tetraethyl (4,4'-diphosphonate-2,2'-bipyridine)

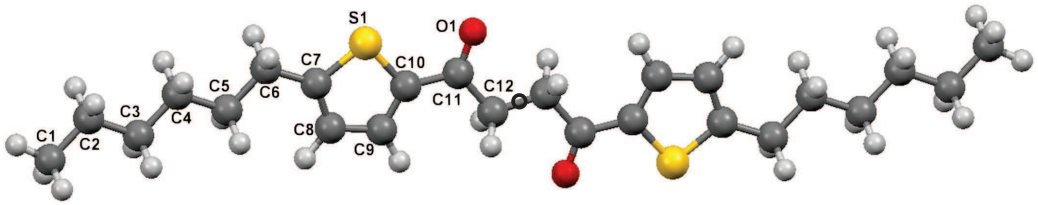


Internal symmetry code: (i) = ***-x, -y, -z***

Experimental and theoretical charge density analysis

Application of a Virtual Atom Model

*1,4-bis (5-hexyl-2-thienyl) butane-1,4-dion**



Triclinic *P*-1

Data collection statistics:

R_{int}	0.053
# Reflections Total	70 042
& Unique	10 227
Absorption Correction	Analytical
T_{min}, T_{max}	0.93, 0.98

Data resolution = 0.48 Å

a = 5.047(1) Å	*α* = 98.79(5)°
b = 6.979(4) Å	*β* = 112.845(4)°
c = 16.333(6) Å	*γ* = 95.39(5)°
Z = 1	*V* = 564.3(7) Å³

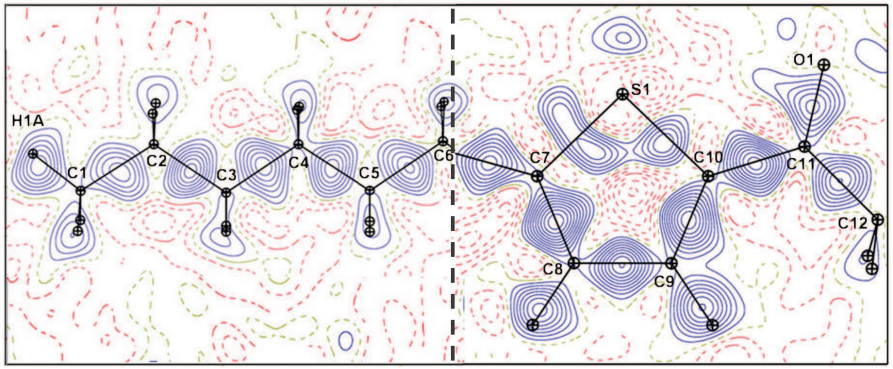
Temperature = 100K

*Ahmed *et al.* (*Acta Cryst. A….In prep.*)

Experimental deformation density after spherical atom model refinement

using Shelxl & MoPro

Contour level: 0.05 e/Å³ 1Å Fourier residual map



$$\Delta\rho_{\exp}(\vec{r}) = {}^{1}\!/_{V} \sum_{H} [K^{-1}|F_o| - |F_{IAM}|] \exp(i\varphi_{IAM}) . \exp(-2\pi i \vec{H} . \vec{r})$$

The diffraction data are very good as:

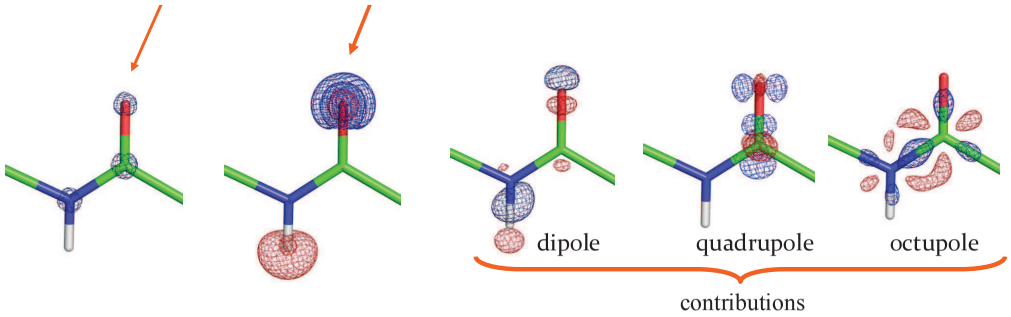
1. Bonding density is very well defined.
2. Lone pairs on O1 and S1 atoms clearly visible.

R(*F*) = 5.3 %

Hansen & Coppens multipolar atom model

$$\rho_{atom}(\mathbf{r}) = \rho_{core}(r) + P_{val}\kappa^3 \rho_{val}(\kappa r) + \sum_{l=0}^{l_{max}} \kappa'^3 R_{nl}(\kappa' r) \sum_{m=0}^{l} P_{lm\pm} y_{lm\pm}(\theta,\phi)$$

Spherical *Non-spherical*



Acta Cryst. A34, 909. (1978)

Number of refinable parameters

1 Scale + Spherical Atom Model

N [3 (xyz) **+6** Uij **]** 9 parameters per atom

\+ Multipolar Atom Model

$N_{non\text{-}H}$ [**1**Pv+ **1**K + **1**K' + **3** (Dip) + **5** (Quad) + **7** (Oct) + **9** (Hex)] +

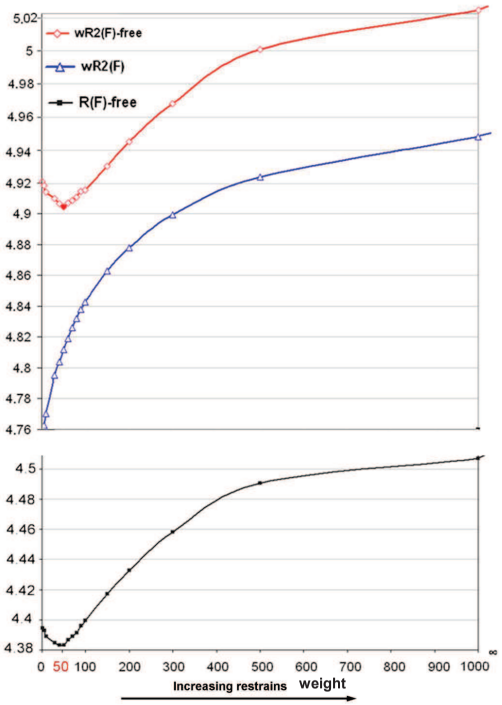
N_{H} [**1** K' + **1** (Dip)] 20-30 parameters in addition

Multipolar atom model has much more parameters to refine,

We may need to use constrains and restraints.

How strong should the restraints be?

R- free refinements*



The diffraction data were divided into 20 sets with 5% of free reflections.

30 different restraints weights were tested for the chemical equivalence *(kappas, valence and multipolar parameters)*

The values of *R*-free factors were averaged. The best restrain was found to be $W_{restrain} = 50$ $(\sigma_{restrain}=0.02)$

*Brünger (1992). *Nature*, **355**, 472-475

*Paul *et al.*, (2011). *Acta Cryst.* B**67**, 365-378.

*Zarychta *et al.*,(2011). *Acta Cryst.* B**67**. 250-260.

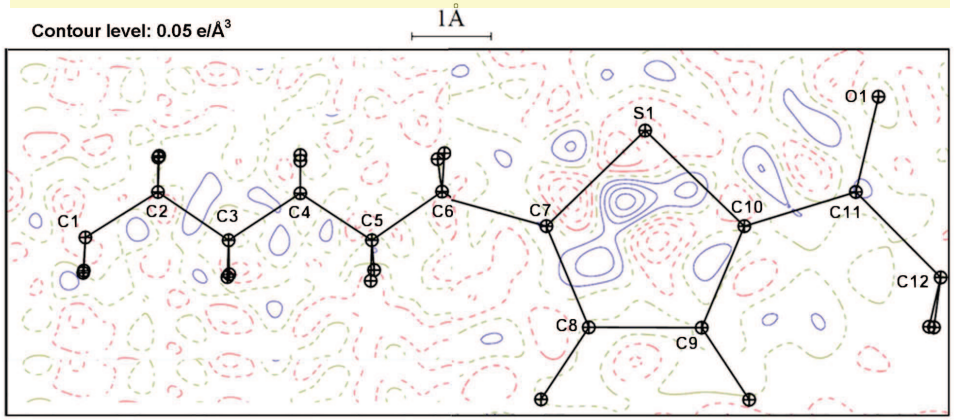
MoPro refinement and residual electron density

using R*free* refinement

***R*(*F*) % final = 4.3** **$wR^2(F)$ % final = 4.9**

C2-C5, C7 & C10 and C8 & C9 were constrained (P_{val}, *K*, *K'* & P_{lm}) together during refinement.

Contour level: 0.05 e/Å^3



$$\Delta\rho_{res}(\vec{r}) = \frac{1}{V}\sum_{H}[K^{-1}|F_o| - |F_{Mult}|]\exp(i\varphi_{Mult})\exp(-2\pi i\vec{H}.\vec{r})$$

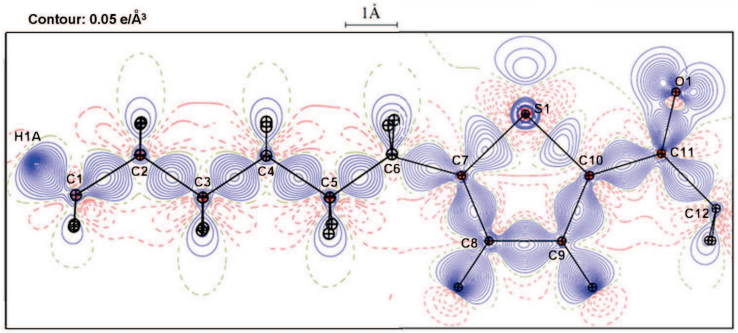
Not much electron density is left to refine. The model is complete.

Static deformation electron density

positive / negative / zero

(a)

Experimental data



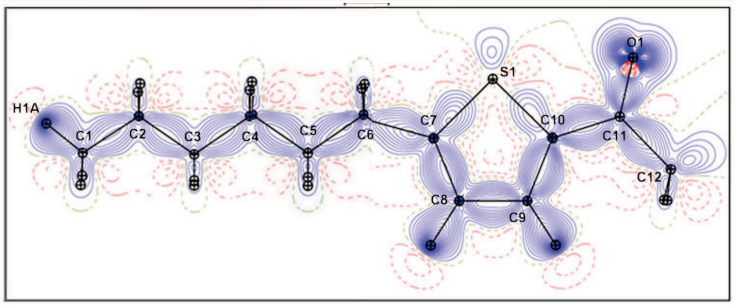
(b)

Theoretical

Structure factors

CRYSTAL09

DFT B3LYP
6-31G**

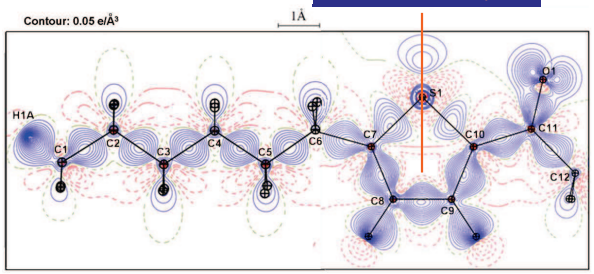


$$\Delta\rho_{stat}(\vec{r}) = \rho(\vec{r}) - \rho_{sph}(r)$$

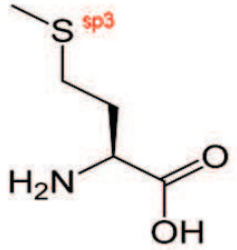
Theoretical work was performed in collaboration with Dr. Ayoub Nassour at CRM2

Difference in lone pair electrons on Sulphur sp^2 & sp^3

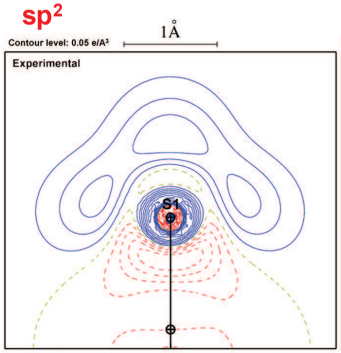
Plane of the lone pairs

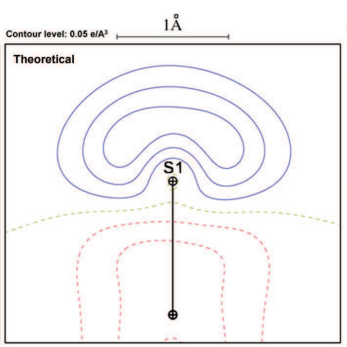


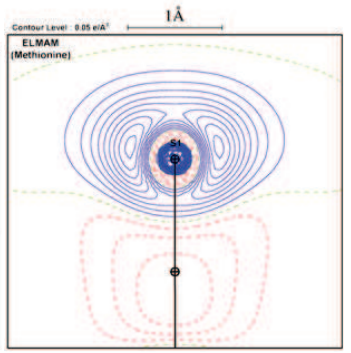
Methionine*



sp^2

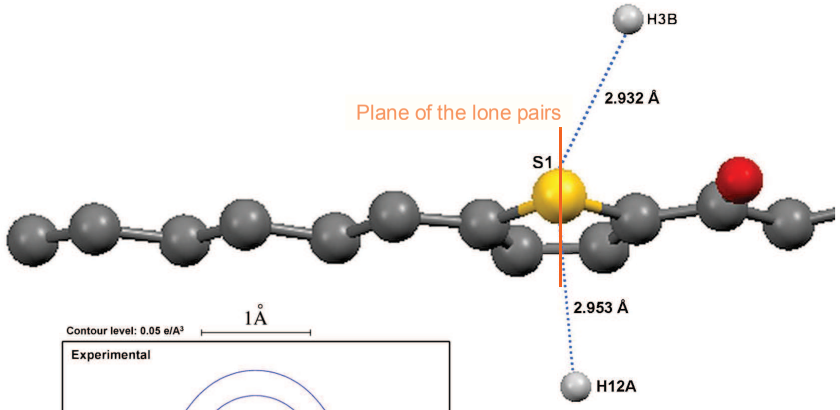


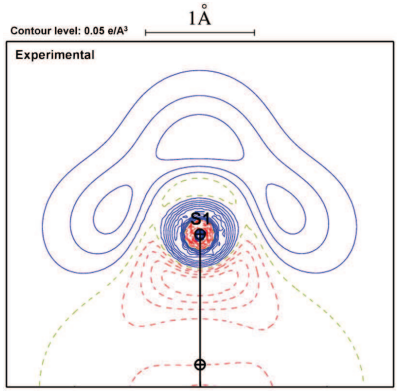




*Guillot, R., *et al*. (2001). *Acta Cryst.* B**57**, 567-578.

Hydrogen bonding with lone pairs on S atom is directional



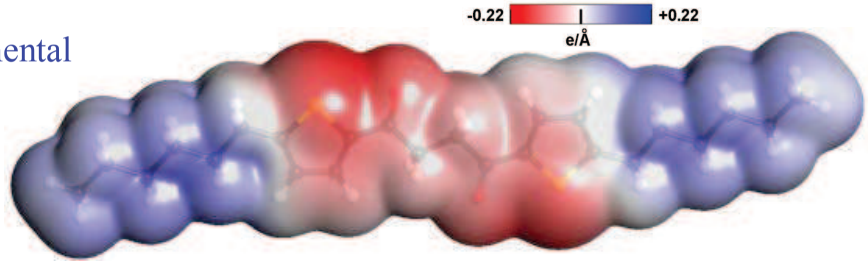


H3B= x+1, y-1, z ; H12A = x-1, y, z

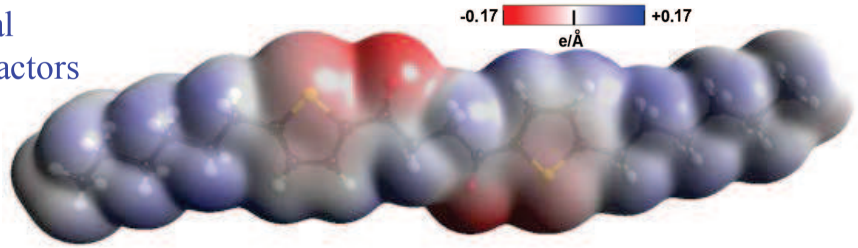
The electrostatic potential

3D electron density surface coloured according to the Electrostatic Potential *Contours @ 0.005 e / Å³*

Experimental



Theoretical
Structure factors



Computed by VMoPro & MoProViewer*

*Guillot, B. (2011). *Acta Cryst.* A67, C511-C512.

Hansen & Coppens multipolar atom model

$$\rho_{atom}(\mathbf{r}) = \underbrace{\rho_{core}(r) + P_{val}\kappa^3\,\rho_{val}(\kappa r)}_{Spherical} + \underbrace{\sum_{l=0}^{l_{max}} \kappa'^3\,R_{nl}(\kappa' r) \sum_{m=0}^{l} P_{lm\pm}\,y_{lm\pm}(\theta,\phi)}_{Non\text{-}spherical}$$

The virtual atom model*

$$\rho(\mathbf{r}) = \sum_{atoms}[\underbrace{\rho_{core}(r) + P_{val}\kappa^3\,\rho_{val}(\kappa r)}_{Spherical}] + \underbrace{\sum_{vir} P_{vir}\kappa_{vir}^3\,\rho_{vir}(\kappa r)}_{Virtual}$$

Strategy for virtual atom model refinement

1- The virtual atoms were initially placed in the middle of two atoms.

2- The initial charges P_{vir} were equal to zero.

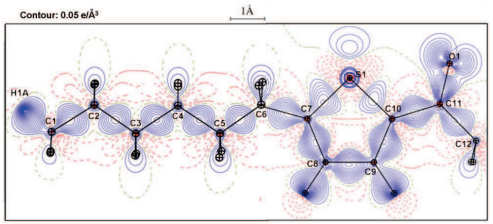
3- The U_{ij} parameters for the H & virtual atoms were constrained.

4- Refinement of P_{vir}, κ and position on bond

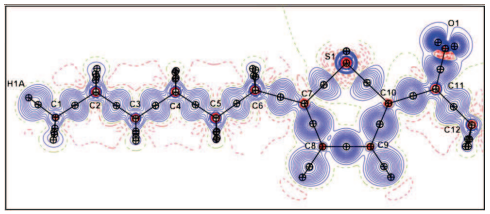
*Dadda *et al.* (2012). Acta Cryst. A

Static deformation electron density

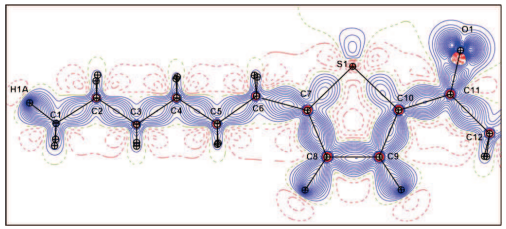
Contours @ 0.05 e / Å³



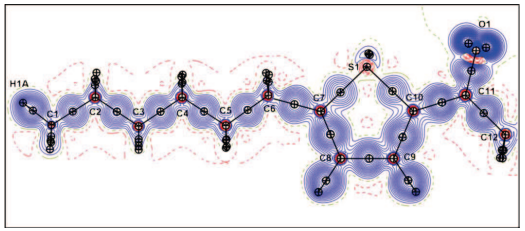
Experimental Multipolar *R (F)* = 4.3%



Experimental Virtual *R (F)* = 4.4 %



Theoretical Multipolar *R (F)* = 0.33%



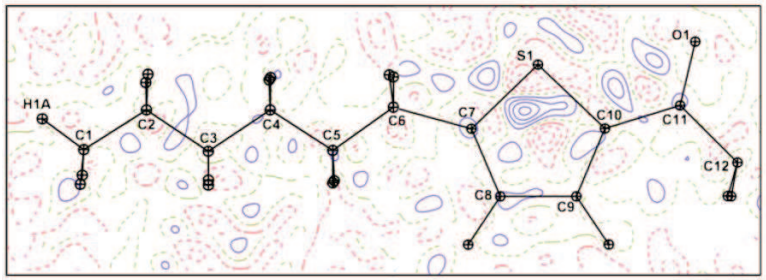
Theoretical Virtual *R (F)* = 0.51 %

Residual electron densities with experimental data

Contours @ 0.05 e / Å³ *Sinθ/λ = 0.7 Å⁻¹*

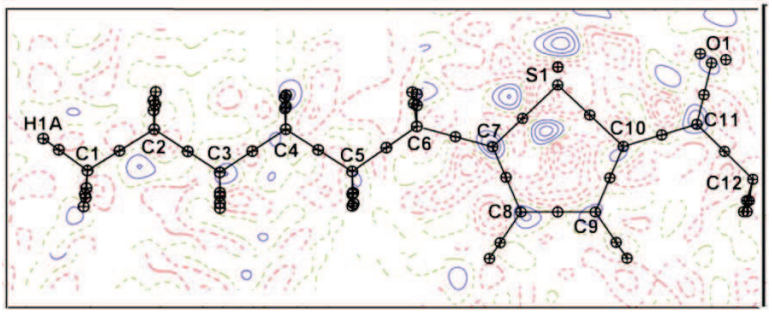
Multipolar model

R (F) = 4.3%



Virtual atom model

R (F) = 4.4 %



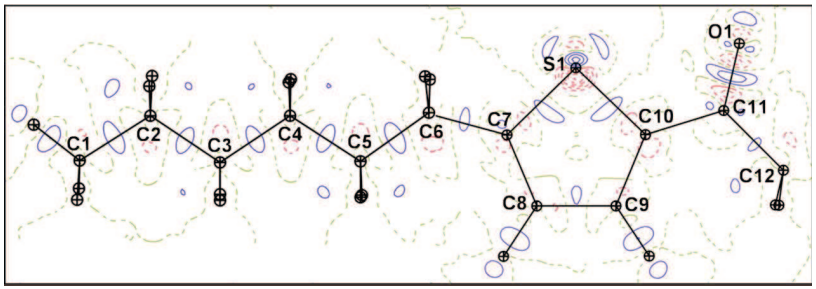
The two models behave similarly for experimental data

Residual electron densities with theoretical data

Contours @ +/- 0.05 e / Å³ sin *θ/λ* = 1.2Å⁻¹

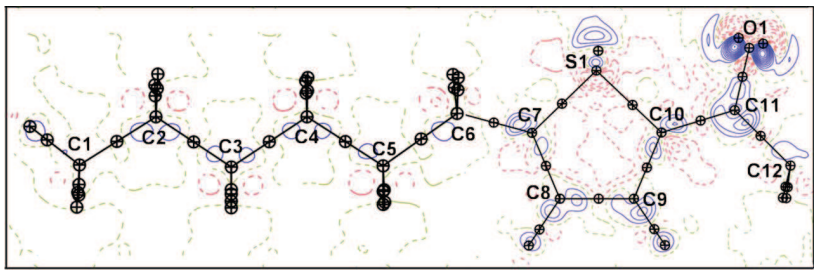
Theoretical Multipolar

R (F) = 0.33%



Theoretical Virtual

R (F) = 0.51 %



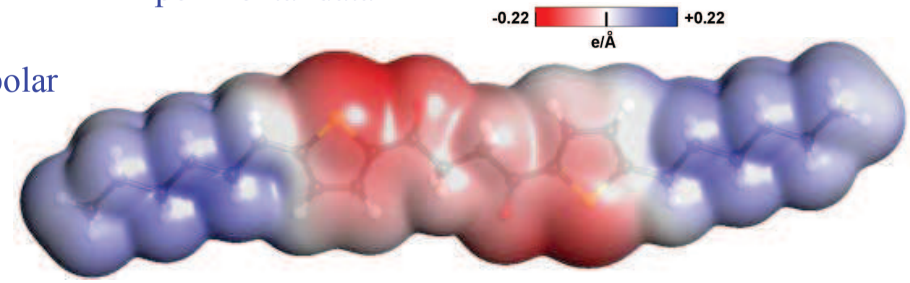
The virtual atom model behaves well except for lone pairs in case of theoretical data

Electrostatic Potential

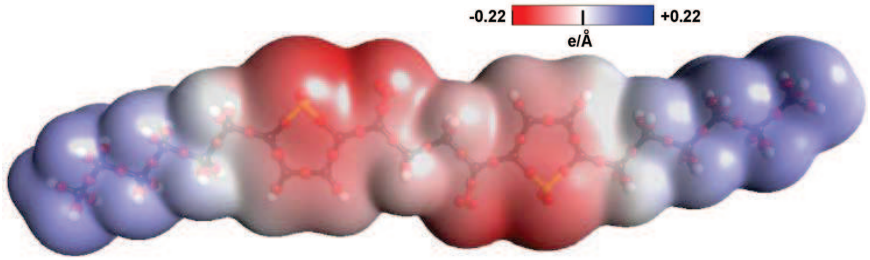
Experimental data

Contours @ 0.005 e / Å^3

Multipolar



Virtual



OUTLINE

1. Structural disorder in Thiophenes

2. Charge density analysis of a Thiophene compound

3. Lone pairs electron density and H bonds

4. Transferability of electron density parameters

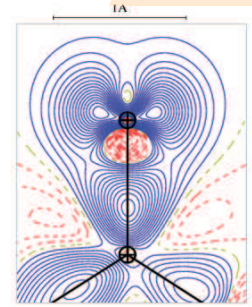
   a - Small molecule

   b - Protein

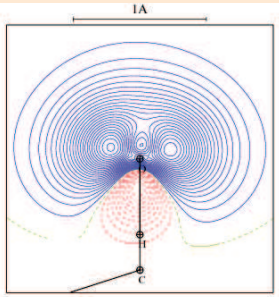
5.Conclusions & Perspectives

Lone pairs electron density of Oxygen atom

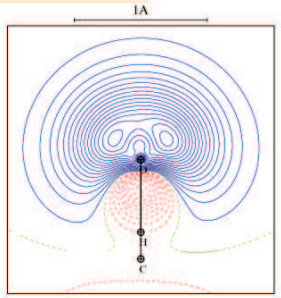
DFT



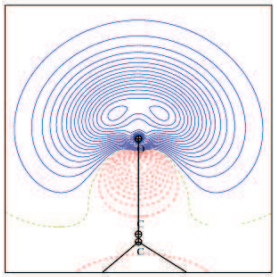
Carbonyl



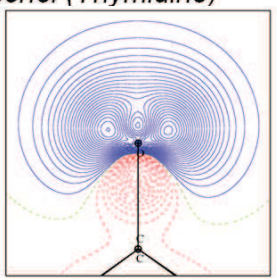
Alcohol (Thymidine)



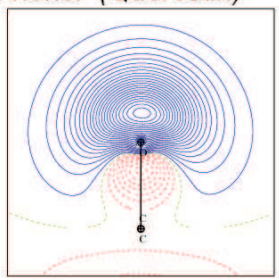
Phenol (Quercetin)



Ester (coumarin)



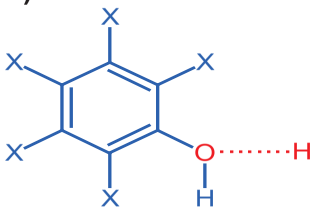
Dimethyl Ether



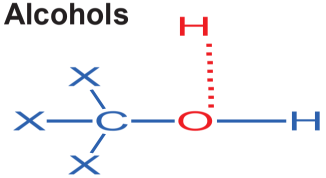
Cyclic Ether (Quercetin)

Some oxygen acceptors

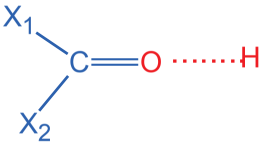
1) Phenols



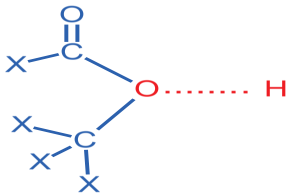
2) Alcohols



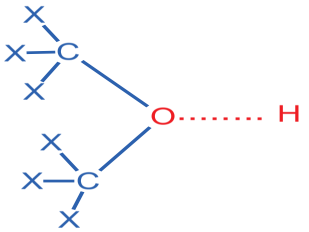
3) Carbonyl



4) Esters



5) Ethers

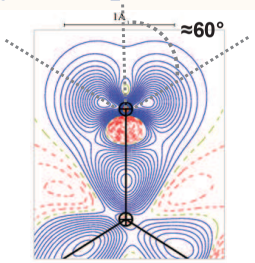


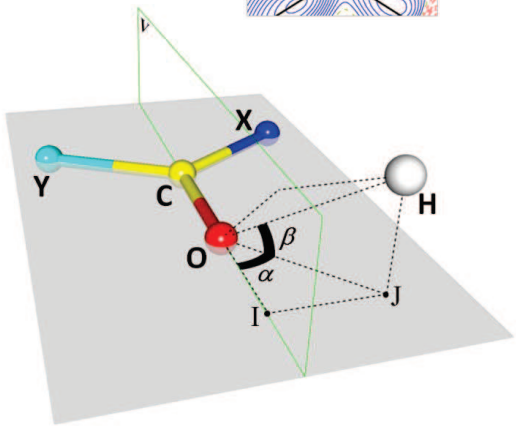
Number of structures retrieved

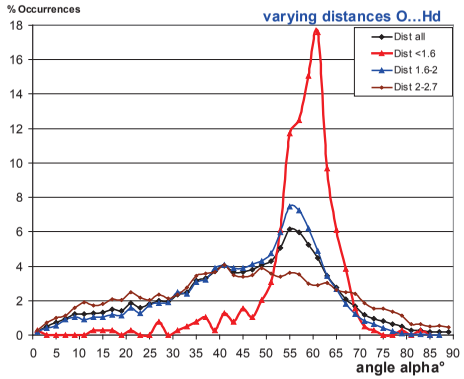
Cambridge Crystallographic Data Centre

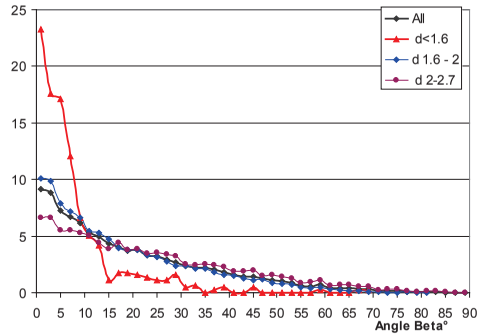
Acceptor \ Donor	-O-H	-N-H
Carbonyl	18 645	12 332
Phenols	1 287	401
Alcohols	8 840	1 484
Esters	176	88
Ethers	1 739	628

Stereochemistry of H bonds with Carbonyl acceptors

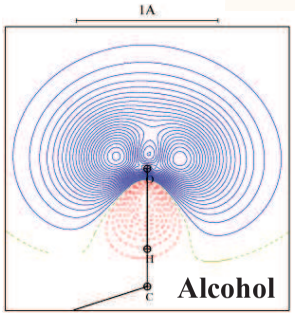


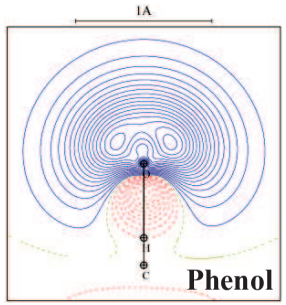


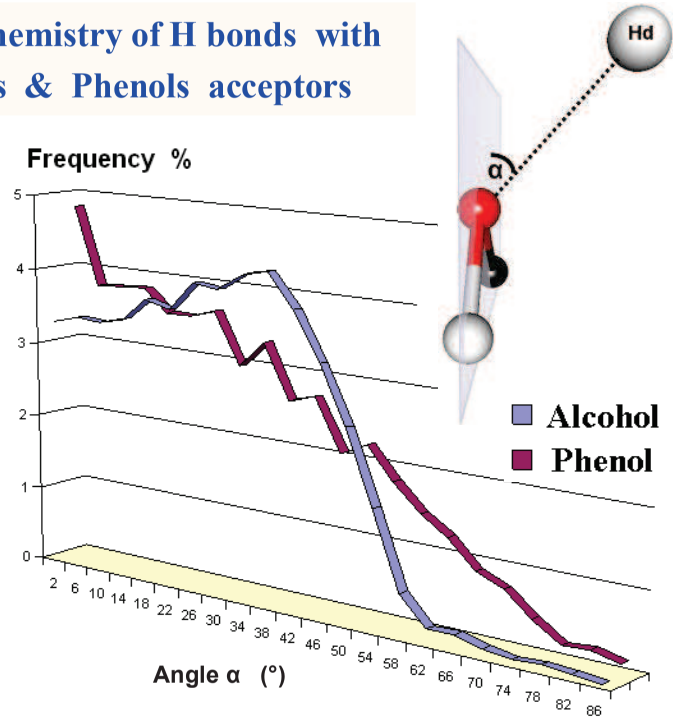




Stereochemistry of H bonds with Alcohols & Phenols acceptors







In alcohols, there is a tendency for angle α to be larger than in phenols

OUTLINE

1. Structural disorder in Thiophene

2. Experimental charge density analysis

3. Lone pairs electron density and H bonds

4. Transferability of electron density parameters

a - Small molecule

b - Protein

5.Conclusions & Perspectives

The ELMAM2 Library

Experimental Library of Multipolar Atom Model

Small unit cells precise charge density studies

Atom types definition :

Based on chemical species, hybridization state, first and second neighbor atoms.

Averaging of **atom types** parameters.
Only **significant multipoles** are kept.
Using **local Symmetry** of deformation electron density

ELMAM2

Transfer ELMAM2 **to**
-large systems
-small molecules
-lower resolution data

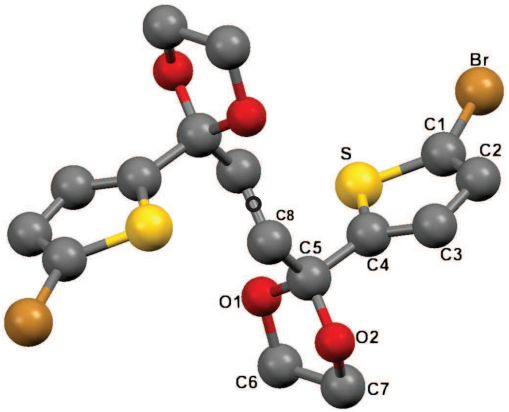
Used as :
- Starting set of parameters
- Properties computation

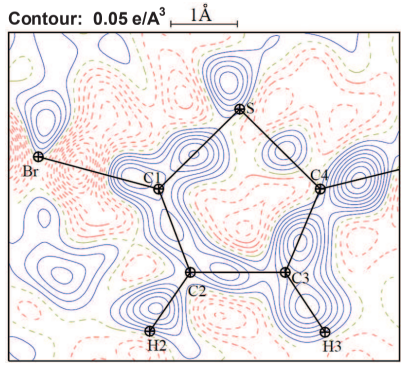
Domagala *et al.* (2011) *Acta.Cryst.* A

ELMAM2 database transfer

2, 2'-(Ethane-1, 2-diyl) bis [2-bromothiophen-5-yl)-1, 3-Dioxolane*

C16 H16 Br2 O4 S2





Experimental Deformation Density

Fourier map after IAM refinement (spherical atoms)

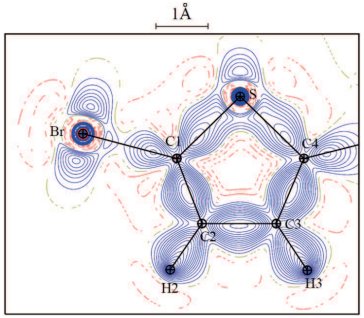
Data resolution = 0.83 Å T=100K

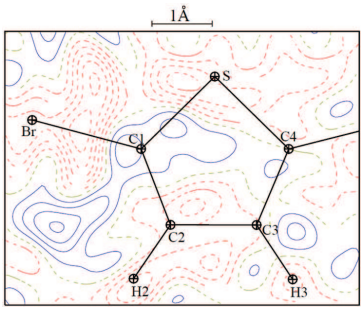
Monoclinic C2/c

a = 19.255(1) Å	β = 112.85(4)°
b = 5.7800(4) Å	V = 1736.6(2) Å^3
c = 16.933(6) Å	Z= 4

*Ahmed *et al.* (2011). *Acta Cryst.* C67, o329-o333.

ELMAM2 database transfer





Residual electron density map after ELMAM2 transferred

$\rho_{atom}(\mathbf{r}) = \rho_{cor}(r) + P_{val}\,\kappa^3\,\rho(\kappa r) + \Sigma\,\kappa'^3\,R(\kappa' r)\,\Sigma\,P_{lm\pm}\,y_{lm\pm}$

The parameters of P_{val}, κ & κ' and $P_{lm\pm}$ were transferred from ELMAM2.

Br atom type provided by S. Dahaoui of CRM2.

In the subsequent refinement, only *Scale*, *XYZ* and U_{ij} *parameters* were refined.

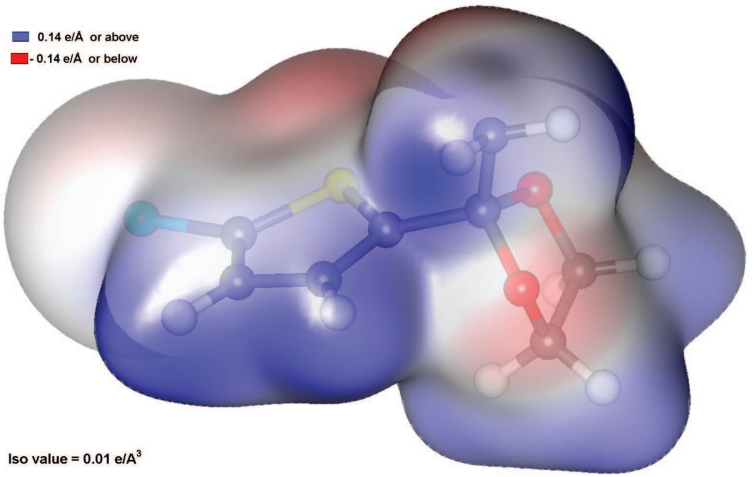
Refinement statistics

	IAM	Multipolar Atom Model
$R(F)$ %	2.7	2.2
$wR^2(F)$ %	4.1	3.1

Not much electron density is left to refine, hence the model is complete.

Electrostatic potential of asymmetric unit*

Computed using ELMAM2 database transfer



Electron density

*Computed by VMoPro & MoProViewer

Applications of

ultra high resolution methods

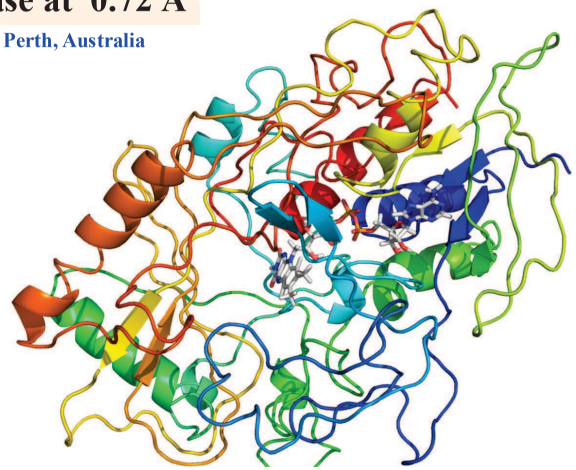
to Proteins

Protein Cholesterol Oxidase at 0.72 Å

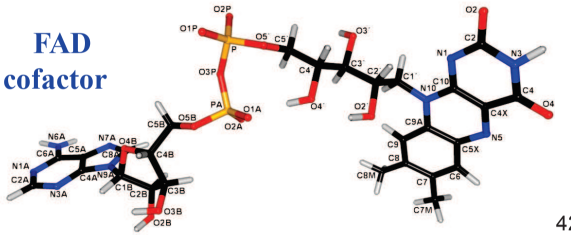
In colloboration with Prof. Alice Vrielink, Perth, Australia

Organism	Brevibacterium *sterolicum*
Molecular weight	55 kDa
Method	X-rays diffraction
Space group	P 2_1
Cell dimensions	a = 51.273A°
	b = 72.964 A°
	c = 63.036A°
	β =1 05.18°
# Measured reflections	465 385
# Unique reflections	444 511
$R(F)$ %	11.5
R free (F) %	14.0
$wR^2(F)$ %	13.5
# of Residues	**499**
# Residues in α-helices	125
# Residues in β-sheets	53

The structure was refined with *Phenix



FAD cofactor

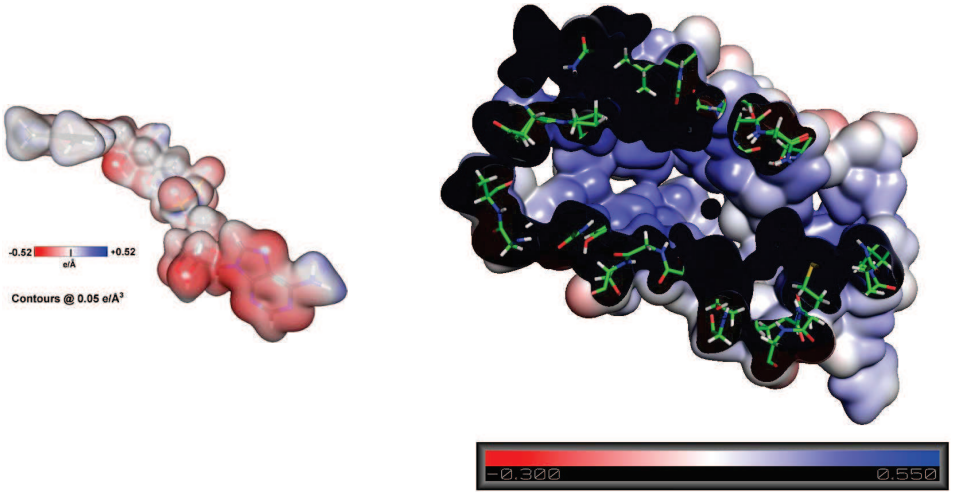


Flavin adenine dinucleotide

* Afonine *et al.*, Acta Cryst D, 2012

The electrostatic potential of the active site cavity*

Computed using ELMAM2 database transfer

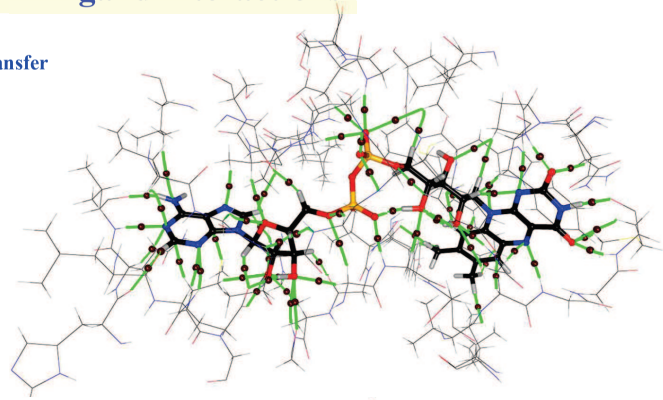


*Computed by VMoPro & MoProViewer

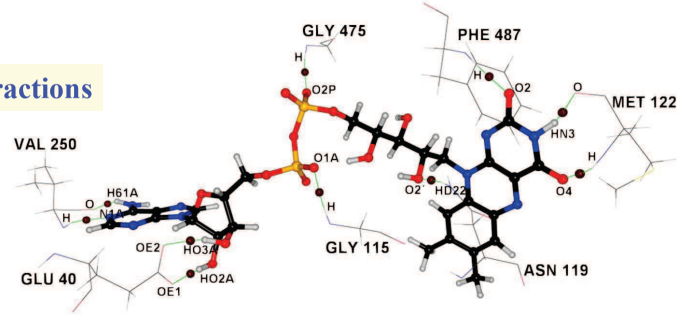
Images are made with PyMol and MoProViewer.

Topology of Protein-Ligand interactions

using ELMAM2 database transfer



Shortest interactions



Intra-molecular interactions in FAD

O1A···HO4'

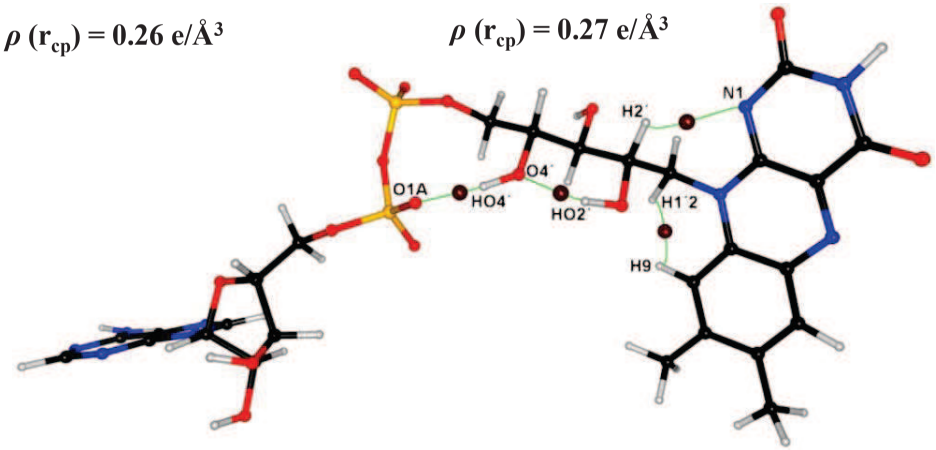
d = 1.79Å

ρ (r_{cp}) = 0.26 e/Å^3

O4'···HO2'

d = 1.78Å

ρ (r_{cp}) = 0.27 e/Å^3



These intra-molecular interactions are supposedly responsible for the specific geometry of the FAD molecule.

OUTLINE

1. Structural disorder in Thiophenes
2. Charge density analysis of a Thiophene compound
3. Lone pairs electron density and H bonds
4. Transferability of electron density parameters
   - a - Small molecule
   - b - Protein

5. Conclusions & Perspectives

Conclusions

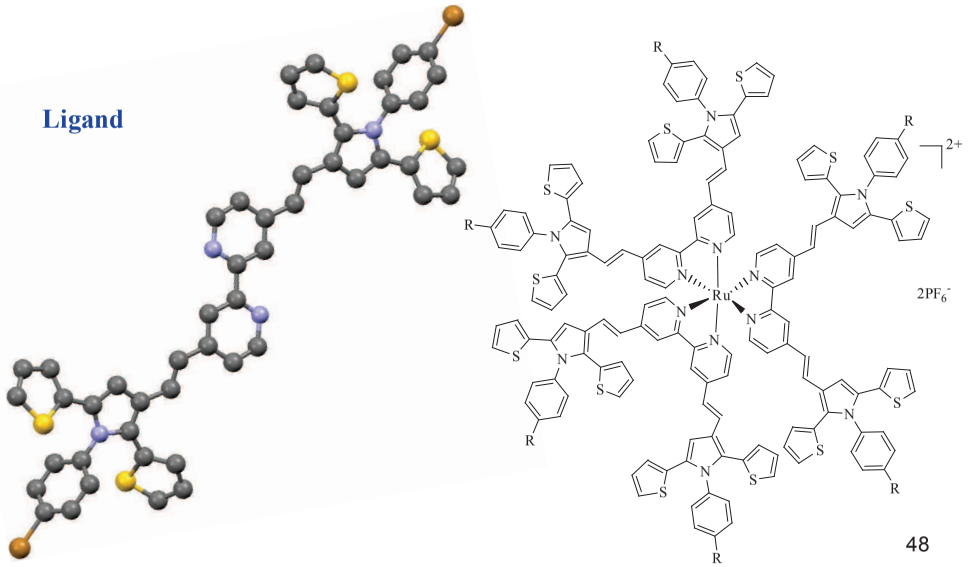
1. Challenging X-rays structures of thiophenes were solved which helped to understand the phenomenon of disorder.
2. Using ultra-high resolution crystallographic methods, experimental and theoretical charge density analysis of molecules which have significance for organic solar cells was carried out.
3. An extensive use of crystallographic database (CCDC) was made to find how lone pairs electron density affects H bonding.
4. In the absence of a high resolution data, the 'transferability principle' was successfully applied to a small molecule and a large protein to calculate electron derived properties.

Perspectives

1. Structures upto ligand have been correctly determined.

   Structures of complexes will be studied next.

Ligand

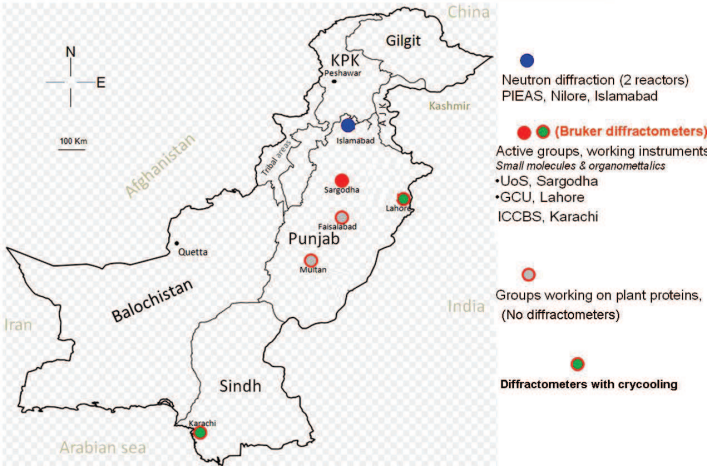


Perspectives

2. Introduction of modern crystallography in Pakistan.

Current activity is limited to structure reports only.

Crystallographic centers in Pakistan



Pakistan is a major partner in the SESAME synchrotron being constructed in Jordan.

http://www.sesame.org.jo/sesame/

Acknowledgements

Supervisors:

Prof. Claude Lecomte,

Dr. Christian Jelsch &

Dr. Benoit Guillot

Collaborators:

Sajida Noureen , Prof. M. Beley & Dr. P. Gros,

(*SRSMC, SOR group, Nancy*)

Prof. Alice Vrielink, (*Perth, Australia*)

All colleagues at CRM2

My family in Pakistan

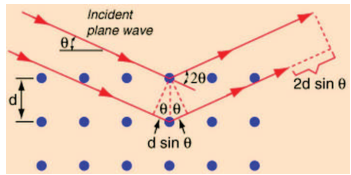
The Higher Education Commission of Pakistan (HEC) for financing this PhD thesis. (www.hec.gov.pk)

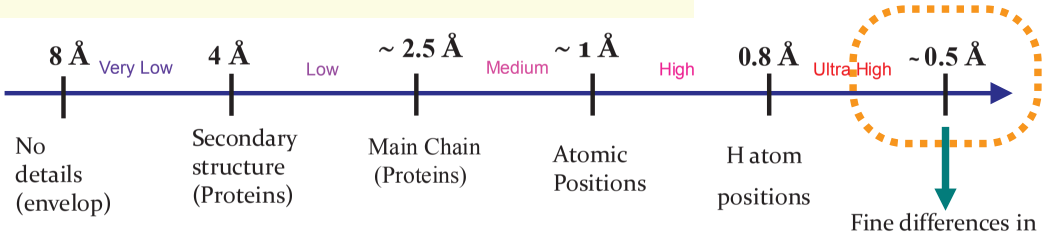
Thank you!

Resolution in Crystallography

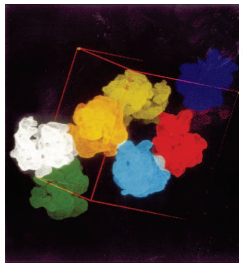
Bragg's Law

$n\lambda = 2d \sin\theta \longrightarrow d = n\lambda / 2\sin\theta$ (Å)

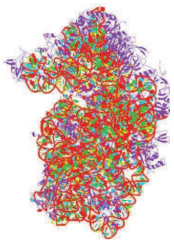




Ribosomes(70S), 10Å

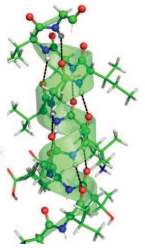


Younath *et al*, 1998



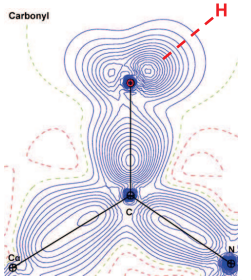
2XZM

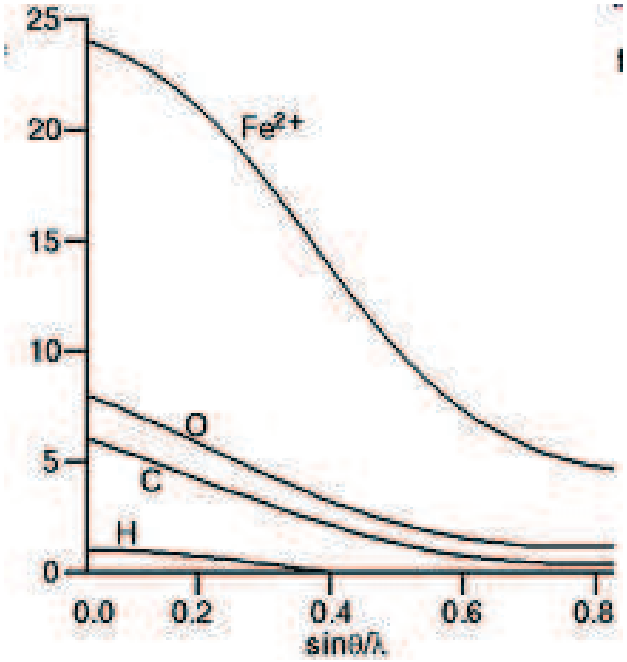
3.9 Å



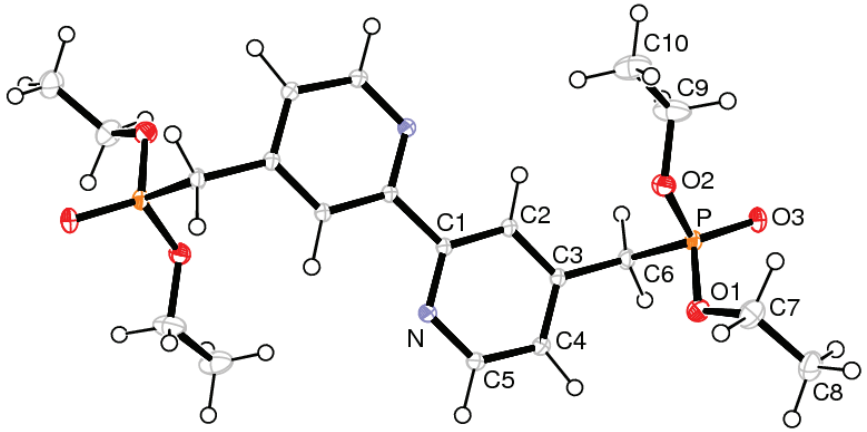
3DO3

2.5 Å

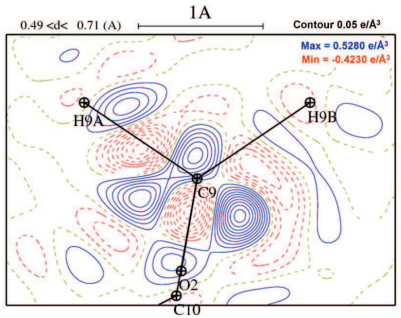


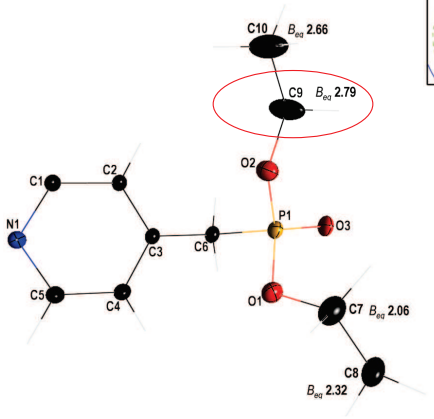


VII Charge density analysis

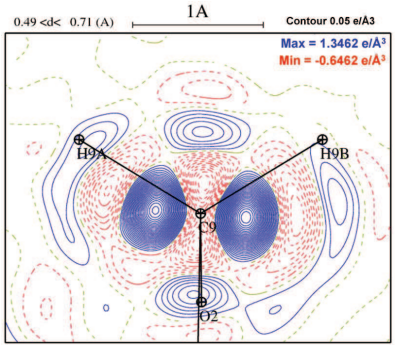


(1)- Simple Multipoles refinement

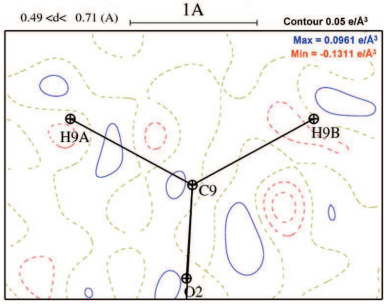


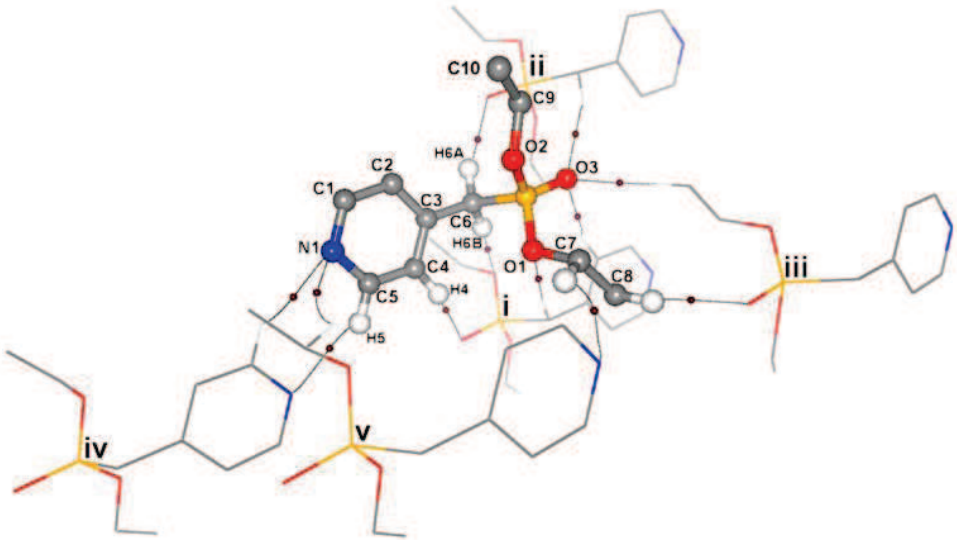


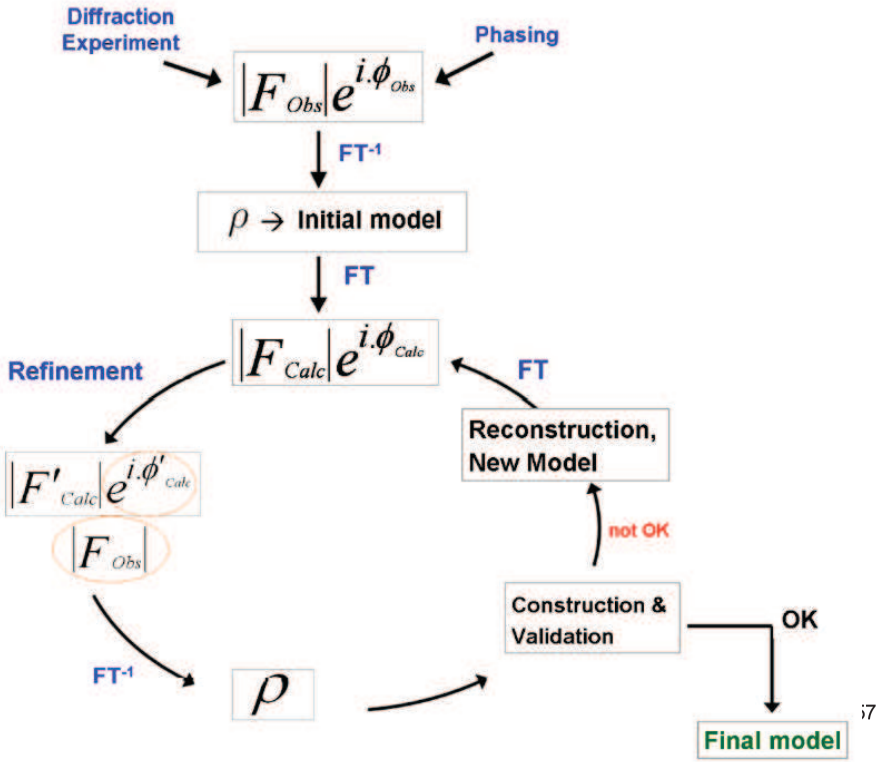
(2) - 3rd order Gram Charlier anharmonic



(3) – 4th order Gram Charlier anharmonic

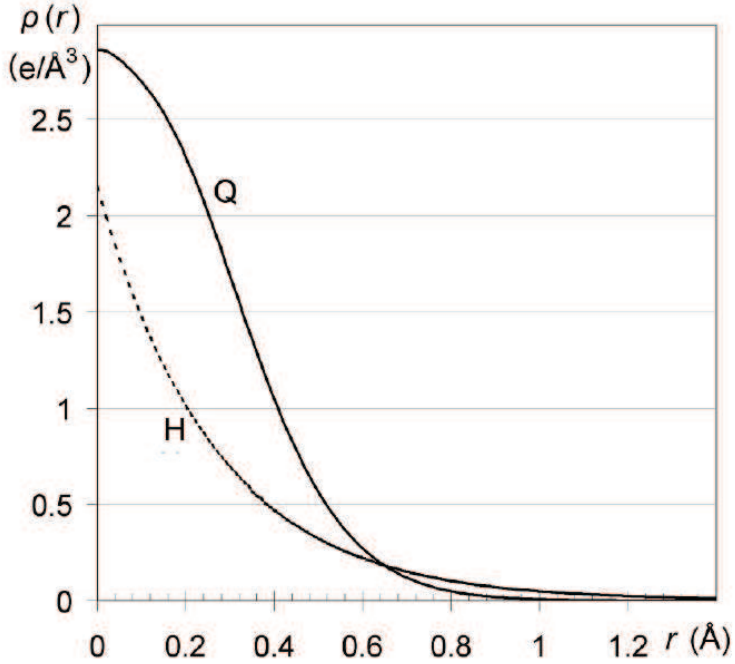






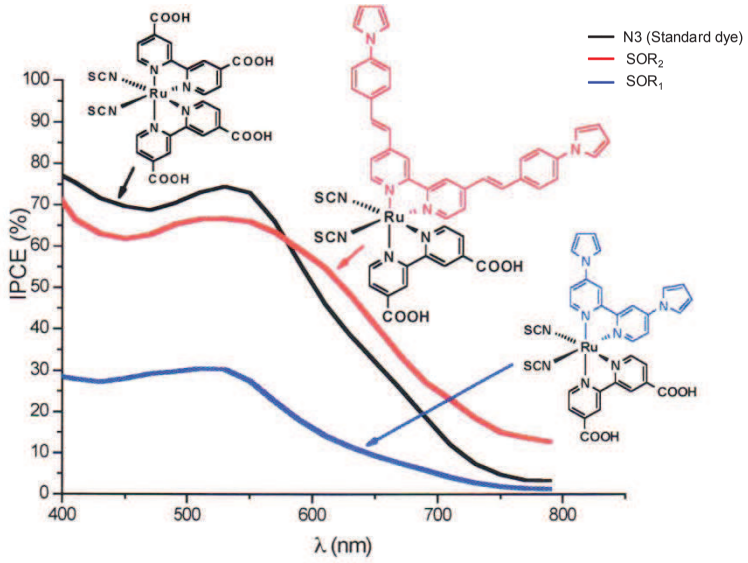
Electron Density of Virtual Atoms

Electron density of a spherical virtual atom Q with P_{vir}=1 as a function of the distance r to the nucleus. For comparison, the same curve is shown for a hydrogen atom.



Previous Results of SOR Group

Thiophene result !!!!!!



Nazeeruddin *et al. J. Am. Chem. Soc.* 1993, **115**, 6382.
Grabulosa, A.; Martineau, D.; Beley, M.; Gros, P.C.; Cazzanti, S.; Caramori, S.; Bignozzi, C.A. *Dalton Trans.* 2009, 63.