RECENT RESULTS OF THE LABORATORY OF CRYSTALLOGRAPHY IN BORDEAUX

Michel HOSPITAL

Laboratory of Crystallography. University Bordeaux I

The Laboratory of Crystallography is a part of University Bordeaux I which is one of the 3 Universities of Bordeaux located on the campus of Talence, a suburb situated south west of Bordeaux. The University Bordeaux I has two main components: Sciences and Laws-Economics with 18.000 students, about 600 professors or lecturers and 50 laboratories. For the scientific part there are 400 researchers and engineers from the CNRS (Centre National de la Recherche Scientifique) distributed in 30 laboratories associated to the CNRS.

The laboratory of Crystallography is one of them, associated under the number: UA 144 CNRS. It belongs to both the Physics and the Chemistry departments of the CNRS and to the Physics Unity of the University.

1. Composition of the laboratory

- Permanent Scientists: 28 (with 12 lecturers and professors)
- Engineers and Technicians: 12
- Postgraduate students: 18
- Post-doc and visitors: 5

2. Equipment of the laboratory

X-ray analysis of single crystals

- Weissemberg and precession cameras,
- 2 Nonius CAD 4 diffractometers with N_2 liquid equipment,
- Nonius rotating anode generator,

– Huber-Inel diffractometer for low temperature (15K – 300 K) measurements,
– High pressure equipment (up to 20Kb) data collection.

X-ray analysis of polycrystals
– Siemens D500 high resolution diffractometer,
– 1D curve counter Inel for kinetics and phase transitions,
– Guinier-Simon and Guinier-Lenné apparatuses.

Energetics equipment
– Thermal analysis – AED Dupond de Nemours,
– Thermal analysis – DSC Perkin-Elmer,
– Microscope with thermosystem.

Computing
– Main computer SUN 330,
– Micro computers SUN, IBM, ATARI, HP, DIGITAL, etc.
– 4 Silicon Graphics stations,
– Internal network: Ethernet working with Unix system,
– Connetion to the network of the University and National Centers,
– All the main software for X-ray analysis of molecular and macromolecular compounds and molecular dynamics.

Chemistry
– Peptide synthesis: Applied Biosystems,
– Oligonucleotides: Applied Biosystems,
– HPLC: Waters and Varian
– All the equipment for organic compound synthesis.

3. Domains of research

A simple look at the equipment list easily allows one to see that our main activity concerns the X-ray analysis of molecular and macromolecular compounds such as single crystals and polycrystals.

Till June 1991 the activities of the laboratory could be presented as seven scientific operations.

1. Structures and physicochemical properties,
2. Interactions DNA - drugs,
3. Interactions protein - drugs,
4. Low dimensionality materials,
5. Molecular composites - polymers

6. Molecular alloys and energy stocking,
7. Diffusion defects in crystals.

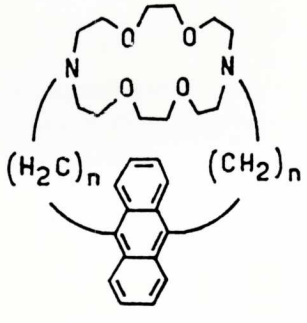
4. Results

To present some examples of our results I took into account the presentation of N.B. Chanh on "Molecular composites" and of T. Granier on "Structural study of transition spin compounds" which give examples of operations 4 and 5.

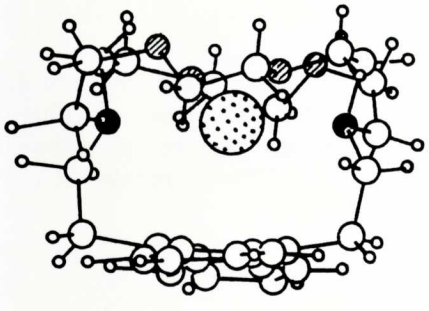
I selected examples of other operations.

4.1. Anthraceno-cryptands with Ag+ (Operation n° 1)

The study of exciting spectra and fluorescence emission spectra in solution or in solid state presents unusual interactions between anthracene and cations such as Ag^+ or Tl^+. Anthraceno cryptands are very well adapted to impose the contact between Ag^+ and anthracene and in this condition to induce a charge transfer complex. Our proceeding was to make such complexes and to study them in UV fluorescence and to perform their X-ray analysis.



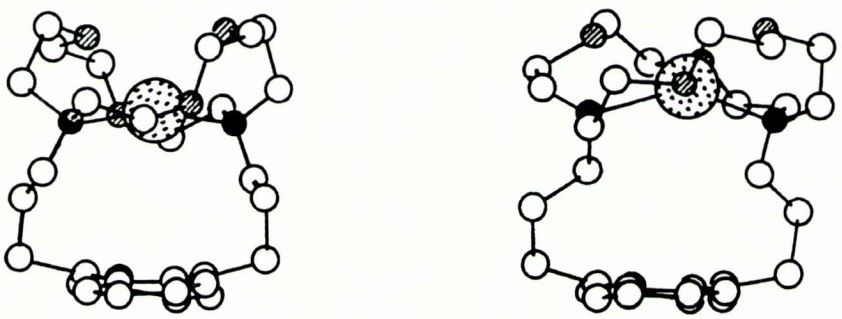
A_{nn} n = 2,3



A_{22}/Ag^+

The fluorescence of cryptands is very sensitive to the solvent and to the cations. In solution there is a monomeric and excimeric fluorescence emission explained by an interaction between N and the aromatic ring. This interaction is weakened in presence of cations like Na^+ or K^+ and we observe an increase of the intensity of monomeric emission with a decrease of the exciplex form. The complex A_{22}/Ag^+ presents the same spectra.

In the crystal we observe a full extinction of the fluorescence possibly explained by a very short distance between Ag^+ and the aromatic ring. X-ray analysis confirms this explanation by showing a short distance Ag^+/plane of aromatic ring (3.09 Å) very close to the addition of the Van der Waals radii of a carbon atom and Ag^+ ion (1.26 Å).



A_{33}/Ag^+

The X-ray analysis of the complex A_{33}/Ag^+ presents two conformations. The two distances Ag^+/anthracene ring are longer than for A_{22} (4.50 and 5.10 Å). The fluorescence of the complex is observed in solution and in the solid state. This result involves a new formation of exciplex Ag^+/aromatic ring.

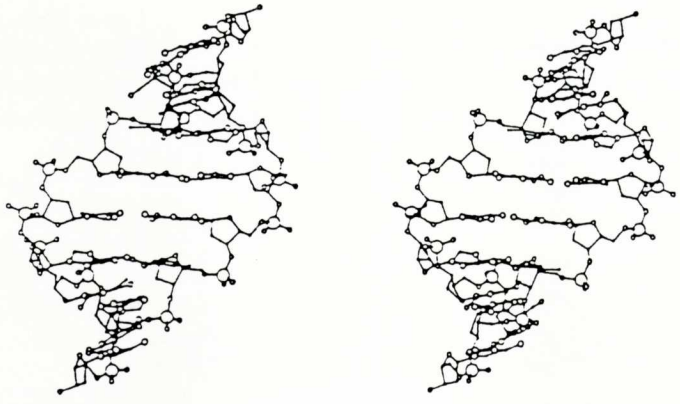
According to the length of the "arms" of the cryptands we are able to prepare Ag^+ complexes with charge transfer in a fundamental state without luminescence (A_{22}) or in the exciplex state with luminescence (A_{33}).

Responsible for the work: Y. Barrans and P. Marsau.

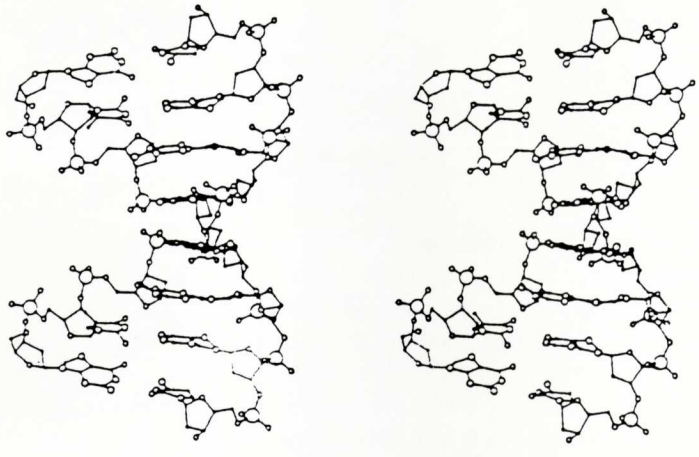
4.2. *DNA fragments*

Several years ago we started a structural study of antitumoral drugs. Some of them have the property to intercalate DNA specifically between the C-G base pairs and we solved the structure of the complex r(CpG) Celiptium. We synthesized and purified several oligonucleotides having a C-G base pair in the sequence. The self complementary sequence d(GTACGTAC)$_2$ crystallized in two different crystalline forms.

The first form belongs to the $P4_32_12$ space group with one octanucleotide single strand and 46 water molecules in the asymmetric unit.



(a)

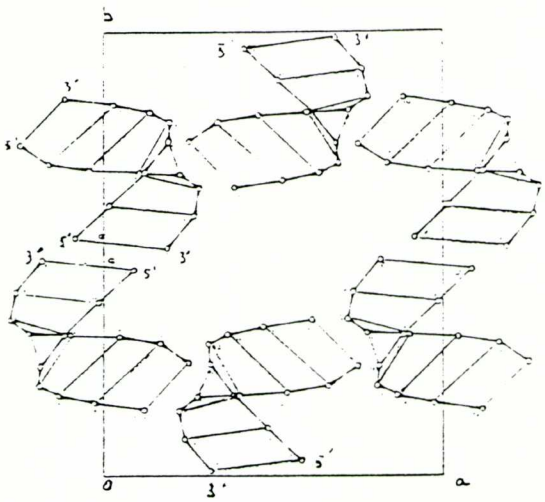


(b)

Stereo representation of the structure of d(GTACGTAC)$_2$. Carbon, nitrogen, oxygen and phosphorous are drawn with different radii. (a) View into the minor groove along the dyad symmetry axis. (b) View perpendicular to the minor groove.

The resolution of the data is 2.3 Å, the structure was solved by molecular replacement and the refinement was carried out with NUCLSQ (Konnert-Hendrickson method) and XPLOR (energy minimization) programs. All the computing work was performed on a 3D graphic station which is absolutely necessary to follow the progress of refinement in the reciprocal space and the direct space at the same time.

The second form belongs to the $P2_12_12$ space group with one double strand and 91 water molecules in the asymmetric unit. The resolution is 2.3 Å and the structure was solved with molecular replacement too. Refinement and positions of water molecules were made as in the previous structure.



Both structures show that $d(GTACGTAC)_2$ belongs to the A type DNA with a general C3' endo conformation for the sugars, an anti-configuration for the glycosidic bonds and a gauche– gauche+ configuration for the α and γ bonds of the phosphate sugar chain. Base pairs display usual Watson-Crick hydrogen bonds.

In the orthorhombic form the structure is close to the A form described in the fiber studies. It is interesting to note that the stacking of terminal bases homologous with the binary axis creates a pseudo oligonucleotide with 16 base pairs as shown in the figure above.

We synthetized and purified several other sequences but up to now we did not obtain crystals of good enough quality to start a structural study either of the oligonucleotides themselves or of complexes with antitumoral drugs.

Responsible for the work: C. Courseille, A. Dautant, B. d'Estaintot.

4.3. *Gramicidin A*

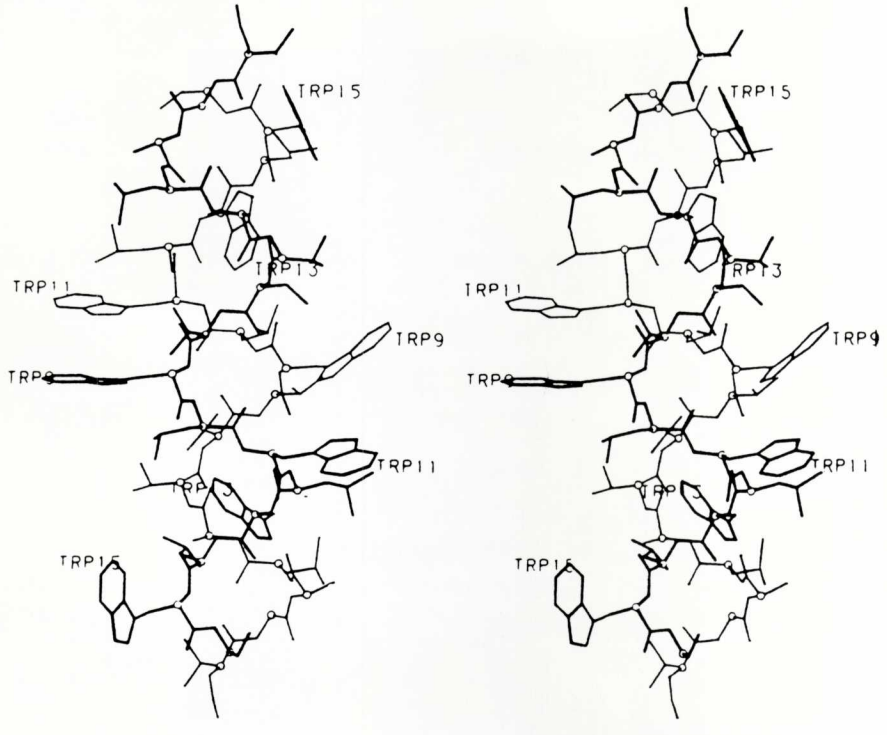
The structure of the gramicidin A was solved in collaboration with researchers of the Medical Foundation of Buffalo. Crystals are monoclinic P_{21}

and obtained from methanol solution. The resolution of the data is 1.5 Å. The structure was solved by a complex way using molecular replacement and rotation function. Refinement was made with FRODO and PROLSQ programs. After the location of 30 amino acids (213 atoms) it was possible to identify only 9 molecules of methanol.

The main characteristics of the structure are a β helix dimer with an antiparallel form, and a position of Trp9 and Trp11 residues approximately perpendicular to the helix axis.

A comparison with the orthorhombic structure already solved in Buffalo shows a different diameter for the helix (5.09 Å instead of 4.8 Å) and some differences in the position of residues.

Responsible for the work: C. Courseille, G. Precigoux, M. Hospital, D. Langs* and G.D. Smith* (* Medical Foundation of Buffalo).



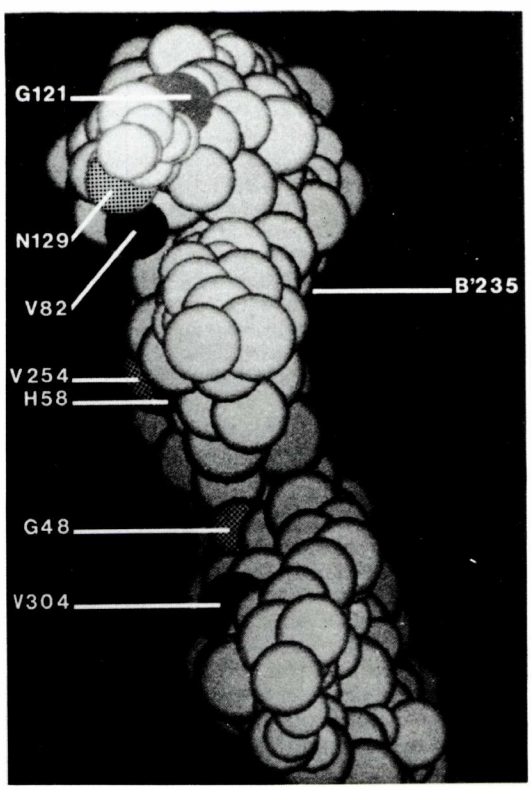
Stereoview of the monoclinic form of the $\uparrow\downarrow\beta^{5,6}$ helix dimer of gramicidin A crystallized from methanol. The two strands of the dimer are differentiated by light and heavy lines, and only the major conformers of the disordered sidechain residues are illustrated. Note that the orientation of the indole rings of Trp-9 and Trp-11 on both strands is approximately normal to the helix axis.

4.4. *Structural analysis of retrovirus proteins*

To present our recent work about retroviruses I have to say that we work on two proteins: envelop glycoprotein (gp 120) and protease of two retroviruses: HTLV-1 (Human T cell Leukemia Virus) and BLV (Bovin Leukemia Virus).

– *Glycoprotein of HTLV-1 envelop : gp120*

For several years we developed a scientific programme to approach the prediction of three-dimensional structure of a protein with only the knowledge of the primary structure (sequence of aminoacids). Starting with a prediction of the secondary structure with the use of statistic data and the hydrophobic/hydrophilic residue distribution, and then setting up the folding potential features in 3D structures of known proteins, it is possible to propose a model for the unknown protein.



Model of gp120. Each sphere represents one residue.

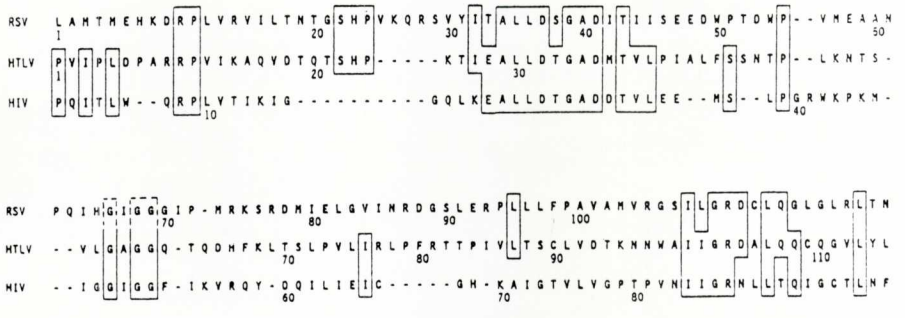
We proposed such a model for the envelop protein of HTLV-1 virus built from the hemaglutinine of the influenza virus. This model is very useful to propose some specific antibodies corresponding to the external part of the molecule. Up to now all the antibodies obtained with the epitope present on the surface of the model were found. Setting up a neutralizing antibody could lead to a specific vaccine.

– HTLV-1 protease (115 amino-acids)

The only biological thing to know is that a protease is necessary for the maturation of the virus in an infected cell, so if we are able to inhibit the protease of a virus, we get the opportunity to avoid the proliferation of the virus and to stop the disease.

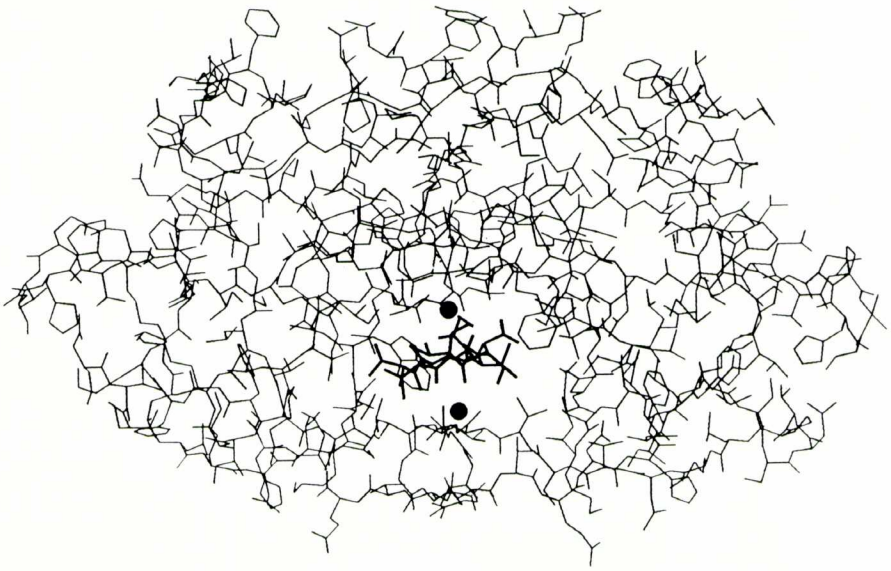
We built a model with the help of the X-ray structures of two other virus proteases HIV-1 (Human Immunodeficiency Virus) responsible for AIDS and RSV (Rous Sarcoma Virus) responsible for chicken bone tumors.

Three-dimensional information allowed us to improve the sequence alignments as shown in the figure below where the conserved residues are contoured.



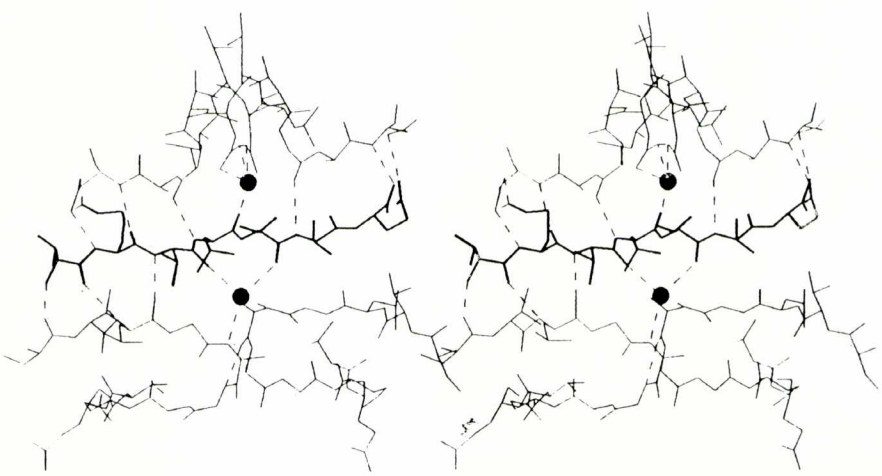
Our model was built using two structures: for the main part with RSV results, and for the flap with HIV results. The protease belongs to the aspartyl protease family and is active in a dimeric form. It is mainly constituted of β sheets with a small α helix and the two monomers are associated by hydrogen bonds at the N and C extremities. In the neighbourhood of the active site, we found a particular β sheet, called "flap", which is involved in the stability of substrates in the site after complexation.

The model was built with the MMS program. Deletions are situated on the outside part of the loops, and have no influence on the general conformation of the molecule. Several substrates were set in the active site with respect to hydrogen bonds between peptide and protease.



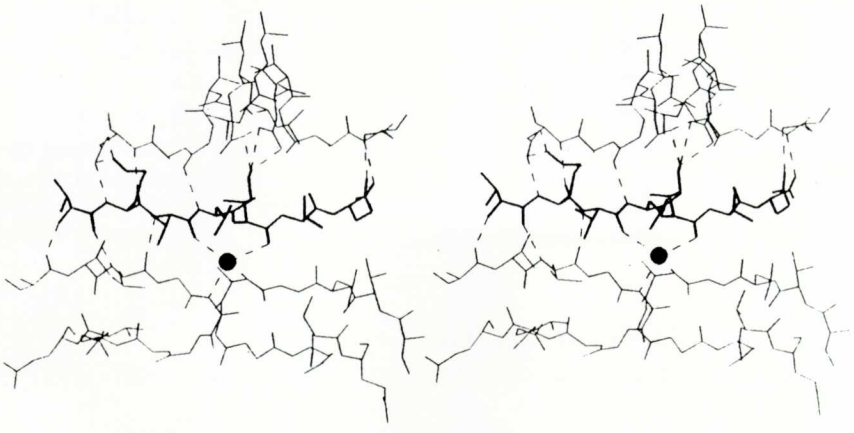
Model of HTLV-1 protease.

All the models of proteases were refined by energy minimization with the XPLOR program. It was possible to locate two water molecules playing a role in the stabilization of the peptides in the active site. The same work with an anhibitor specially synthesized for this protease showed that one water molecule was eliminated.



Stereoview of the active site with a substrate Thr - Lys - Val - Leu - Val - Val - Gln.

All the substrates and inhibitors were synthesized and purified in the laboratory. X-ray analyses of some of them were made to get a more precise description of the conformation. Small crystals of protease were obtained and the work to set up a better inhibitor is in progress.



Stereoview of the active site with an inhibitor.

Responsible for the work: B. Busetta, S. Geoffe, S. Llido, P. Picard, G. Précigoux.

4.5. *Molecular alloys*

In this field we are concerned with both the fundamental knowledge and the industrial applicability of molecular alloys (i.e. molecular mixed crystals).

– Some years ago this scientific topic got an official European framework through the creation of the REALM (Rèseau Européen sur les Alliages Moléculaires) which unites five research groups of three different nationalities (the name of each national manager is italisized):

France: Bordeaux I University (group of Dr. *Y. Haget*)

The Netherlands: Utrecht University (group of Pr. *H.A.I. Oonk*)

Spain: University of Barcelona (group of Pr *M.A. Cuevas Diarte*); Autonomous University (Barcelona) (group of Pr. *E. Estop*); Polytechnical University (Barcelona) (group of Pr. *J. Muntasell*)

General coordinator: Y. Haget.

– The multidisciplinarity approach has enormously increased the insight in the relation between molecular and crystalline structure and macroscopic (thermodynamic) properties. Extensive studies of miscibilities allowed us to analyze the competitivity between two essential factors which govern the syncrystallization: the differences of both shape and size of molecules A and B (for a binary $A_{1-x}B_x$ example) and their differences in nature. Criteria have been established to answer questions such as: what are the necessary conditions for the observation of total or partial miscibility? In other words, how may the limits of molecular alloy formation be foreseen? One of these criteria is the degree of crystalline isomorphism of A and B. Doing that, the necessity to reactualize the concept of isomorphism itself appeared. I chose to develop here these particular results because their validity is very general even if the best way to progress in this field was the opportunity of molecular syncrystallization studies. Two main cases have to be considered.

First case: Total miscibility (Fig. 1)

Each domain of total miscibility corresponds to the possibility to obtain $A_{1-x}B_x$ alloy whatever x is. Each crystalline parameter $p_i(x)$ varies continuously from $x = 0$ to $x = 1$. Only one G curve is required to account for the stability of all the alloys from A to B. All these properties define the isomorphism of A and B.

Second case: Partial miscibility

Miscibility may only be partial either of the eutectic or of the peritectic type. For example, for a given temperature T, there are three domains: two domains of miscibility which go from A to S_A and from S_B to B, separated by a domain of immiscibility where the two limit solid solutions S_A and S_B coexist.

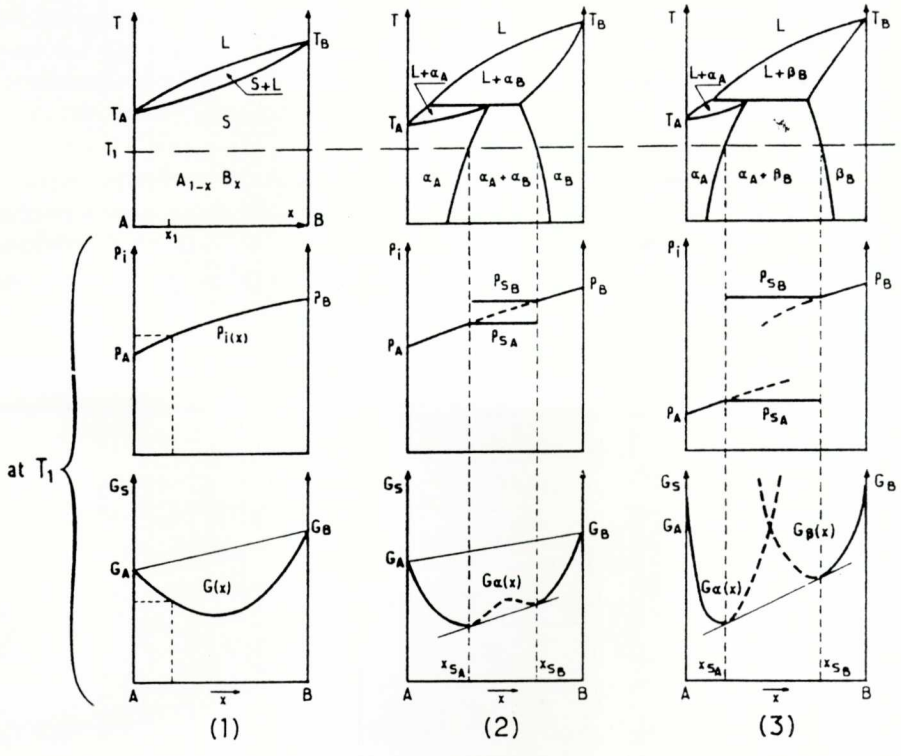
Two cases have to be considered:

i – (Fig. 2). Crystalline parameters continuously vary from $x = 0$ to x_{SA}, and from x_{SB} to $x = 1$; a discontinuity exists between x_{SA} and x_{SB} parameters but it is not very important. Moreover, the stability of all the alloys may be also described here by using only one G curve but this curve presents two inflexion points. The equilibrium conditions are fulfilled by the double tangent line, the mole fractions of the existing phases being the abscissae of the points of contact: x_{SA} and x_{SB}. Under these conditions A and B have to be considered as isomorphic but the degree of isomorphism is not sufficient to ensure a complete miscibility.

ii – (Fig. 3). Here again crystalline parameters continuously vary from $x = 0$ to x_{SA} and from x_{SB} to $x = 1$ but the gap between the parameters of the two limits S_A and S_B is very large. Moreover, one G curve is not sufficient to

account for the stability of all the alloys. Two curves G_A and G_B are needed. The composition of the coexisting phases S_A and S_B are the abscissae of the points of contact of the common tangent line.

In this case, A and B are non isomorphic.



– Concerning the valorization of our materials, we showed that if A and B are judisciously chosen one can generate Molecular Alloys Phase Change Materials (MAPCM) suitable for thermal energy storage. This new class of materials is interesting for various reasons, the main one being that they provide an answer for a lot of temperature ranges where classical PCM are scarce and even non-existent, and above all that one can modify their concentration in order to get *the* material which stores and removes heat at *the* required temperature.

The diversity of MAPCM applications and their feasibility have been proved with prototypes and industrials have already concretely expressed their interest for the "must" brought by these "adaptable materials".

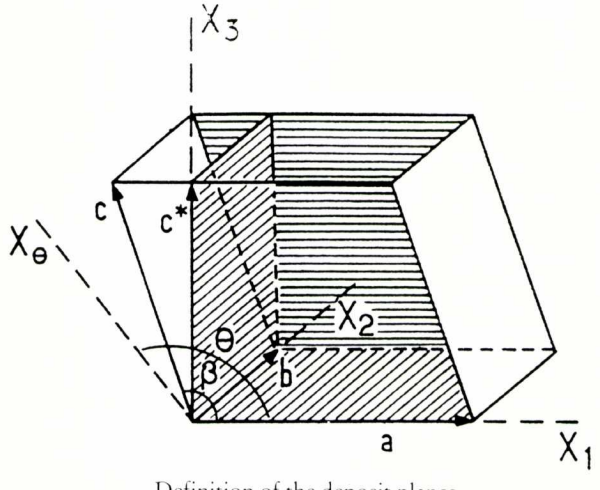
Responsible for the work: Y. Haget, D. Mondieig, V. Soubzmaigne, Ph. Négrier.

4.6. *Diffusion in crystals*

The interpretation of the properties shown by some molecular crystals (electrical conductivity, superconductivity, etc...) needs the knowledge of the structure and also of the typical feature of ponctual and extended defects. Diffusion studies are a very powerful tool to study different defects in molecular crystals. These methods frequently used in metallurgy are less frequent in molecular compounds and if they are performed it is also in one direction which is, in general, the direction perpendicular to the cleavage plane.

We performed a study of the monoclinic crystal of naphtalene (space group P_{2_1}, Z =2). Diffusion tensor is symmetrical. In order to obtain the value of this tensor we measured the diffusion in 4 directions: a, b, c* and x_0.

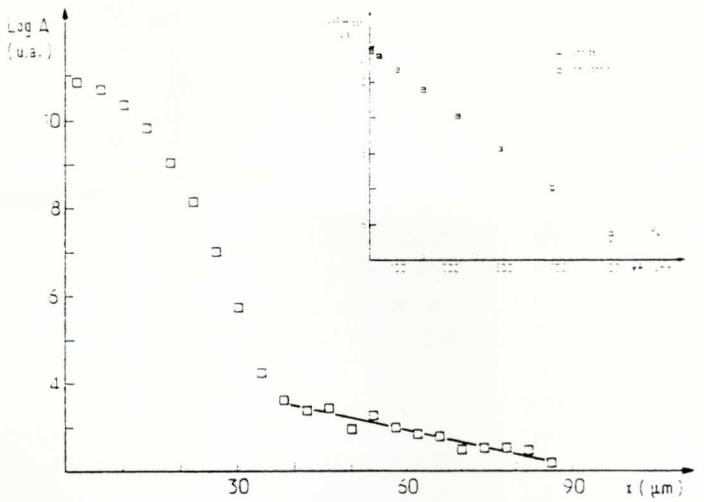
Direction of study	Deposit plane
x_1 direction of $\vec{a}$	(100)*
x_2 direction of $\vec{b}$	(010)
x_3 direction of $\vec{b}$*	(001)
x_0 direction situated in (010) with an angle of 3M/4 with $\vec{a}$	(x_0)*



Definition of the deposit planes.

The principle of the diffusion coefficient measurement in one direction is to deposit a radioactive tracer on a plane perpendicular to this direction and to count the radioactivity of successive thin slices of the crystal. The coefficient D (m^2/s) is obtained with the Fick equation which expresses the concentration as a variable dependent on the time t, the depth x of the slice in the sample and the amount M of deposited compound.

$$c(x, t) = \frac{M}{\sqrt{\pi DT}} \quad \exp \frac{-x^2}{4Dt}$$

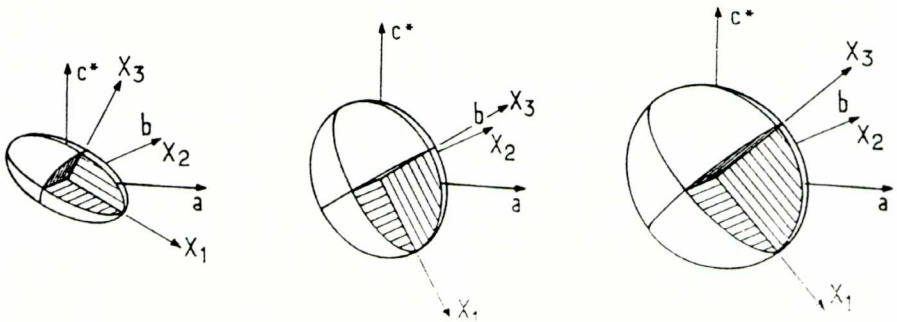


Direction of study x_1, t = 720h, T = 338K

Usually concentration curve shape displays a gaussian profile but we can observe two different areas on the curve obtained for the self diffusion of naphtalene. Until 30 µm we have a gaussian attributed to the diffusion in the volume. Above 30 µm a straight line is obtained and this profile is attributed to diffusion along dislocations. We have to correct the first part of the data by this diffusion using a mathematical treatment (Leclaire et Rabinovitch). So we obtain a corrected diffusion coefficient. For a density of dislocations of about 10^8 m^{-2} we have a coefficient of diffusion in volume of 10^{-17} m^2s^{-1}, and a coefficient of diffusion by dislocation of 10^{-8} m^2s^{-1}.

Such experiments made in 4 directions give self-diffusion tensors and the main results about naphtalene are summarized in the table below as ellipsoids for 3 different temperatures: 338, 343 and 348K.

The symmetry of this property is well indicated with the conservation of the binary axis of the crystal.

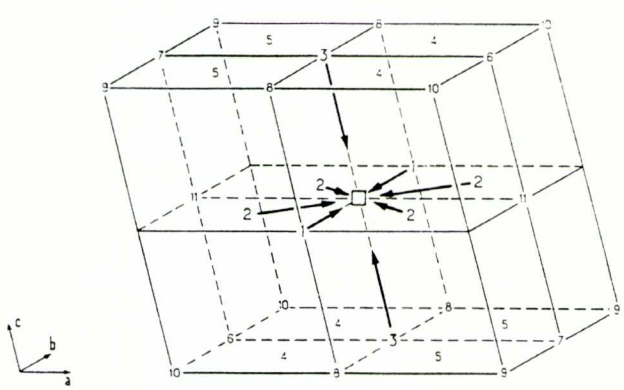


338 K 343 K 348 K

$$\begin{vmatrix} 1.8 & 0 & 0 \\ 0 & 1.7 & 0 \\ 0 & 0 & 0.4 \end{vmatrix} \cdot 10^{-17} m^2 s^{-1} \quad \begin{vmatrix} 2.7 & 0 & 0 \\ 0 & 2.6 & 0 \\ 0 & 0 & 1.6 \end{vmatrix} \cdot 10^{-17} m^2 s^{-1} \quad \begin{vmatrix} 3.8 & 0 & 0 \\ 0 & 2.9 & 0 \\ 0 & 0 & 2.9 \end{vmatrix} \cdot 10^{-17} m^2 s^{-1}$$

A microscopic interpretation of the results could be made in terms of molecular jumps. Considering 8 unit cells with a gap in the center and the symmetry elements of the monoclinic group, we can assume that some jumps are equivalent. For example the jump [1/2 1/2 0] labelled 2 is equivalent to jump [1/2 $\bar{1}$/2 0], [$\bar{1}$/2 $\bar{1}$/2 0] and [$\bar{1}$/2 1/2 0]. We can say the α multiplicity of the jump number 2 is 4.

In the naphtalene structure we have two orientations for molecules and two different types of jumps. The first one associates the initial state and the final state with molecules of the same orientation: it is a pure translation. The second type associates the initial and the final state of different orientation: it is a composition of translation and rotation.



Possible energetically equivalent jumps in naphtalene.

Taking into account all these observations we have to consider only 3 jumps to calculate the energy [010], [1/2 1/2 0] and [001].

q	α_q	l_q(Å)	Movement
1	2	5.98	Translation
2	4	5.10	Translation and rotation
3	2	8.68	Translation

The table above summarizes results for the type of a jump q, multiplicity α_q and the length of the jump and the type of movement. These three jumps are sufficient to explain the volume diffusion of naphtalene.

It was also possible to evaluate the frequency of the different types of jumps and to say that the easiest jumps are those inside the (001) plane.

Responsible for the work: L. Bonpunt, J. Bendani.

SUMMARY

The activity of the Laboratory of Crystallography of the University Bordeaux I is presented with examples chosen among the results obtained during the last two years. The variety of the research subjects ranges from molecular biophysics to crystal physics with the structural studies as the general background.