HAL
open science

Résolution des structures cristallines par diffraction des rayons X et neutrons sur poudres en utilisant les méthodes d'optimisation globale

Luca Palin

To cite this version:

Luca Palin. Résolution des structures cristallines par diffraction des rayons X et neutrons sur poudres en utilisant les méthodes d'optimisation globale. Matériaux. Université Joseph-Fourier - Grenoble I, 2005. Français. NNT : . tel-00077233

HAL Id: tel-00077233

https://theses.hal.science/tel-00077233v1

Submitted on 29 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE JOSEPH FOURIER-GRENOBLE I
SCIENCE & GEOGRAPHIE

THESE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER

Discipline: Physique

Présentée et soutenue publiquement par

Luca PALIN

Le 4 mars 2005

Résolution des structures cristallines par diffraction des rayons X et neutrons sur poudres en utilisant les méthodes d'optimisation globale

COMPOSITION DU JURY:

M. ANNE	Université Joseph Fourier	*Président*
D. LOUËR	Université de Rennes I	*Rapporteurs*
D. VITERBO	Università del Piemonte Orientale "A Avogadro"	
A. N. FITCH	ESRF & Dpt. of Chemistry, Keele University	*Directeur de thèse*

Acknowledgements

I would like to thank the people who have helped me during the course of my thesis.

Andy Fitch and Michela Brunelli for the help during the data collection and the data analysis.

Jon Wright for all his advices in the use of computer programs and the time spent on the beamline ID11 performing the experiments with me.

Phil Pattison for the time he gave me for the experiments performed on the SNBL

Vincent Favre-Nicolin for all the explanations and helps with the program FOX

Alan Coelho for all the explanations and helps with the program TOPAS

Lukas Palatinus for his help with the MEM method.

Daniel Louër and Davide Viterbo for all their advices and corrections made to the initial manuscript

Michel Anne for accepting the position of President of the jury

Mikael Bertolini for his help in the French translation

Contents

Acknowledgements **2**

Contents **4**

Introduction **8**

Introduction (Fr) **10**

Chapter 1 Theory **12**

1.1 Diffraction from a crystal 12

1.1.1 The phase problem 21

1.2 Powder diffraction 26

1.2.1 The Rietveld Method 28

1.2.2 Autoindexing programs 30

1.2.3 Extracting integrated intensities 32

1.2.3.1 The iterative Le Bail algorithm 32

1.2.3.2 The Pawley Method 33

1.2.4 Real space search with Monte Carlo algorithms 34

1.2.4.1 Simulated annealing 35

1.2.4.2 Parallel tempering 39

1.2.4.3 TOPAS and the constrained conjugate gradient routine 42

1.2.5 Fourier map 46

1.2.6 Maximum Entropy Method 48

1.2.6.1 PRIMA algorithm 55

Chapter 2 Instrumentation **58**

2.1 Synchrotron sources 58

2.2 High resolution powder diffraction beamline, ID31/BM16 62
2.3 The Swiss-Norwegian beamlines, SNBL (BM1A) 63
2.4 The materials science beamline, ID11 64

Chapter 3 The low temperature phases of globular organic molecules 66

3.1 Molecular crystals 66
3.1.1 Crystal packing 70
3.2 Plastic crystals 71
3.2.1 7-oxabicyclo[2.2.1]heptane 73
3.2.1.1 Phase I 76
3.2.1.2 Phase II 77
3.2.1.3 Phase III 78
3.2.1.4 Phase IV 85
3.2.2 Norcamphor 90
3.2.2.1 Phase I 92
3.2.2.2 Phase II 92
3.2.2.3 Phase III 98
3.2.3 (1R)-(+)-nopinone 101
3.2.3.1 Phase I 102
3.2.3.2 Phase II 103
3.2.4 (±)2-Azabicyclo[2.2.1]hept-5-en-3-one 107
3.2.5 (+)-endo-2-norborneol 111
3.2.5.1 Phase I 112
3.2.5.2 Phase II 114
3.2.6 (±)-endo-norborneol 117
3.2.7 (±)-exo-norborneol 118
3.2.7.1 Phase I 119
3.2.7.2 Phase II 120
3.2.8 Eucalyptol 121
3.2.9 Un-idexed samples 123

3.2.10 Glassy samples 124
3.2.11 Summary 126

Chapter 4 Neuropeptides 130

4.1 Neuropeptides 130
4.1.1 Opioid peptides 132
4.2 Glycine-Glycine-Phenylalanine 134
4.3 Glycine-Phenylalanine-Leucine 143
4.4 Tyrosine-Glycine-Glycine-Phenylalanine 149
4.5 Leu-enkephalin 149

Chapter 5 Localization of Benzene in a Sodium-X Zeolite 154

5.1 Bare Na-X 156
5.1.1 Cations location 158
5.2 Benzene location 163
5.2.1 Final refinements 167

Chapter 6 Conclusion 180

Conclusion(Fr) 184

Appendix 154

Bibliography 196

Introduction

The primary use of X-ray powder diffraction used to be in identification of compounds, single-phase or multiple-phase materials, *via* their diffraction pattern.

Many solids cannot be grown as single crystals but can be prepared only as microcrystalline powders or, as the case of pharmaceutical compounds, they have to be fully characterised in the form in which they are used. Then if the information that has to be derived from a powder diffractogram is more than simply phase identification and similar to what can be obtained from a single crystal experiment, more sophisticated computational techniques must be developed and used. The change in direction was greatly invigorated by the advent of the Rietveld method in 1969 and the development of advanced radiation sources such as neutron and synchrotron radiation, which brought important improvements in instrumentation, and computational techniques. The Rietveld method has the advantage of avoiding the intermediate step of extracting the individual structure factors or intensities, and therefore the overlapping of reflections did not represent an insurmountable problem.

Before the introduction of the Rietveld method, if a single crystal of a new phase could not be prepared, its structure was likely to remain a mystery. However, even then, if a structural model could be deduced from non-diffraction data, such as chemical analysis, IR/UV-NMR spectroscopy, electron microscopy, the correctness of the model could be checked by comparing its calculated diffraction pattern with the measured one. Using the Rietveld method, not only could a postulated model be verified with the powder pattern, it had also become possible to refine the atomic coordinates. If the compound is completely unknown, in case of high overlap, it is impossible to extract without a structural model the right set of intensities and then Direct Methods used for single crystal structure determination often fail.

Two main improvement allowed the performance of ab-initio structure determination from powder diffraction data: the angular resolution attainable from diffractometers with

synchrotron X-ray sources have led to significant reduction of peak overlap, and the developments in the realm of computing power that are responsible for the fact that computational approaches that would have been unthinkable only 10 years ago have already been put into practice.

In this thesis we have studied three different subjects.

The first is the order-disorder phenomena observed in some globular organic molecular solids. Molecules with near spherical shape often exhibit a high temperature solid phase in which the molecules undergo rapid reorientation. In these circumstances the crystals are easily deformed mechanically and show high-symmetry crystal structures, transforming to more-ordered arrangements as the temperature is lowered. These structures were unknown and their determination has been done performing Monte Carlo simulations.

The second is the opiate family of neuropeptides. These neurotransmitters regulate sensory functions including pain and control of respiration in the central nervous system. The aim of our study was to try to determine the crystal structure of Leu-enkephalin and some of its sub-fragments, just starting form the pure powder in the case that it was already crystalline, or just crystallizing the sample simply using water, without the use of any solvent, in order to be maybe closer to biological reality and see if any structural differences where present, comparing with results from single crystal data found in literature. The determination of the crystal structures has been done performing Monte Carlo simulations.

The third is the location of benzene in a sodium-X zeolite. The zeolite framework was already known and the benzene has been localized by simulated annaeling and by the use of maximum entropy maps.

Introduction (Fr)

L'usage originel de la diffraction de poudre à rayon X consistait en l'identification de composants, mono ou multiphasés, *via* leurs diffractogrammes.

Cette technique est d'autant plus utile que beaucoup de matériaux ne peuvent être préparés sous forme monocristalline alors qu'ils existent sous forme de poudres. De plus, dans le cas des composés pharmaceutiques, ceux-ci doivent être caractérisés dans la forme avec laquelle ils sont utilisés, soit encore une fois une poudre. Alors, si l'on veut tirer d'un diagramme de poudre autant d'informations que d'une expérience basée sur un monocristal, par exemple quelque chose de plus élaboré qu'une identification de phases, des techniques informatiques sophistiquées doivent être développées et utilisées. L'apparition de la méthode de Rietveld, en 1969, ainsi que le développement de sources de radiation avancées telles que les sources neutrons et les radiations synchrotrons, ont induit d'importantes améliorations en matière d'instrumentation et de techniques informatiques. La méthode de Rietveld présente l'avantage d'éviter l'étape intermédiaire d'extraction des intensités des réflexions individuelles et permet en conséquence de surmonter en partie le problème de la superposition des raies.

Avant l'introduction de cette méthode, la structure d'une nouvelle phase ne pouvant être préparée sous forme monocristalline avait de grandes chances de rester inconnue. Cependant, même si un modèle structurel pouvait être déduit de données obtenues par d'autres méthodes comme l'analyse chimique, les spectroscopies IR/UV-NMR, ou la microscopie électronique; la validité du modèle devait être vérifiée en comparant le diagramme de diffraction calculé à celui mesuré. En utilisant la méthode de Rietveld, on peut non seulement vérifier la pertinence d'un modèle donné en le comparant directement avec le diagramme de diffraction expérimental, mais aussi affiner les coordonnées atomiques. Cependant, si la structure du composé est inconnue et le recouvrement des raies important, il devient impossible d'extraire les intensités correctes en l'absence de modèle structurel. C'est pour cette raison que les Méthodes Directes utilisées pour résoudre les structures monocristallines

échouent souvent. Deux grandes avancées techniques ont rendu possible la résolution *ab initio* de structures à partir de diffractogrammes de poudre: d'une part l'amélioration de la résolution angulaire des diffractomètres installés sur les sources de rayonnement synchrotron a conduit à une réduction significative de la superposition des pics, d'autre part l'augmentation spectaculaire de la puissance des ordinateurs autorise l'utilisation courante de méthodes particulièrement consommatrices de temps de calcul, ce que personne ne pouvait se permettre il y a seulement dix ans.

Ce travail de thèse s'articule autour de trois sujets principaux.:

Le premier concerne les phénomènes d'ordre-désordre observés dans certains solides moléculaires organiques de type globulaire. Ces molécules de forme presque sphérique possèdent en effet souvent une phase solide à haute température. Dans ce cas, les cristaux sont facilement déformés mécaniquement, et présentent des structures cristallines de haute symétrie dont l'ordre augmente lorsque la température décroît. Ces structures étaient auparavant inconnues, et nous les avons déterminées en utilisant des simulations de Monte-Carlo.

Le deuxième est centré sur la famille des neuropeptides opiacés**.** Ces neurotransmetteurs régulent, au sein du système nerveux central, les fonctions sensitives telles que la douleur et la respiration. Le but de notre étude était de déterminer la structure cristalline de la Leu-enkephaline, sous sa forme complète ou fragmentée. Pour cela, nous sommes partis d'une poudre pure lorsque l'échantillon était déjà cristallin, ou alors nous l'avons cristallisé en n'utilisant que de l'eau comme solvant, dans le but de se rapprocher de la réalité biologique et de mettre en évidence d'éventuelles différences de structures avec les résultats bibliographiques obtenus à partir de monocristaux. Les structures ont été déterminées grâce à des simulations de Monte-Carlo.

Enfin, le troisième sujet d'étude est la localisation d'une molécule de benzène adsorbée sur des zéolithes X au sodium de structure déjà connue. Ce résultat a été obtenu par simulation et l'utilisation de cartes d'entropie maximale.

Chapter 1 Theory

Ce chapitre débute par un bref exposé de la diffraction de rayonnement X. Nous nous étendrons ensuite plus longuement sur les particularités de la diffraction de poudres et les techniques de calcul utilisées dans l'analyse de données: indexation, extraction des intensités, simulation de Monte-Carlo, affinement de Rietveld, cartographie de Fourier et d'entropie maximale.

1.1 Diffraction from a crystal

In an ideal crystal the molecules or the atoms that constitute it are regularly and periodically disposed in space, this being the configuration of minimal energy. This periodicity can be fully described with the aid of three non-coplanar vectors a, b and c that define a sequence of points that form the crystal lattice. These three vectors define the unit cell, it is possible to find an infinite number of unit cells and we define as primitive those that contain only one lattice point. It can be shown that it is possible to have 14 different Bravais lattices and combining those with the symmetry operators it is possible to have 230 different ways to dispose in space an object periodically repeated. Due to the symmetry, we just need a small part of the unit cell, that is called the asymmetric unit, to describe a crystal structure. This unit is repeated by symmetry to fill the content of the unit cell and the space is filled by the translation vectors that define the lattice.

When X-rays interact with matter, only considering the elastic interactions, electrons start to oscillate with the frequency of the incident beam and they become sources of electromagnetic waves. If we consider only the interaction of the X-rays with an ideal isolated electron density in an asymmetric unit we would have that the diffraction pattern is a continuous function as for the simplest single slit Fraunhofer diffraction case.

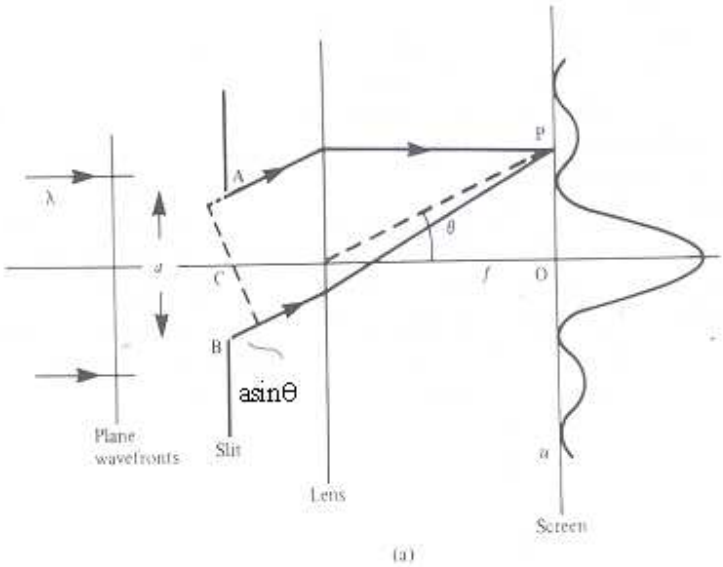


Figure 1-1 Fraunhofer diffraction from a single slit

Now things will be different because in the asymmetric unit in a crystal the electron density is not simply disposed along a line but the kind of reasoning is the same. We see diffraction because the waves originating from different points of the electron density in the asymmetric unit have different distances to cover and this causes a phase difference that causes the diffraction. We can sum all these infinite number of phases and we end up with an integral exactly as in the single slit case.

Due to periodicity in an ideal crystal we have an infinite repetition of asymmetric units and as those are ordered we can observe interference phenomena. The logic behind this is the same as passing form a single-slit to a N-slit-grating. Because in a crystal the number of unit cells is huge the diffraction maxima will be very narrow.

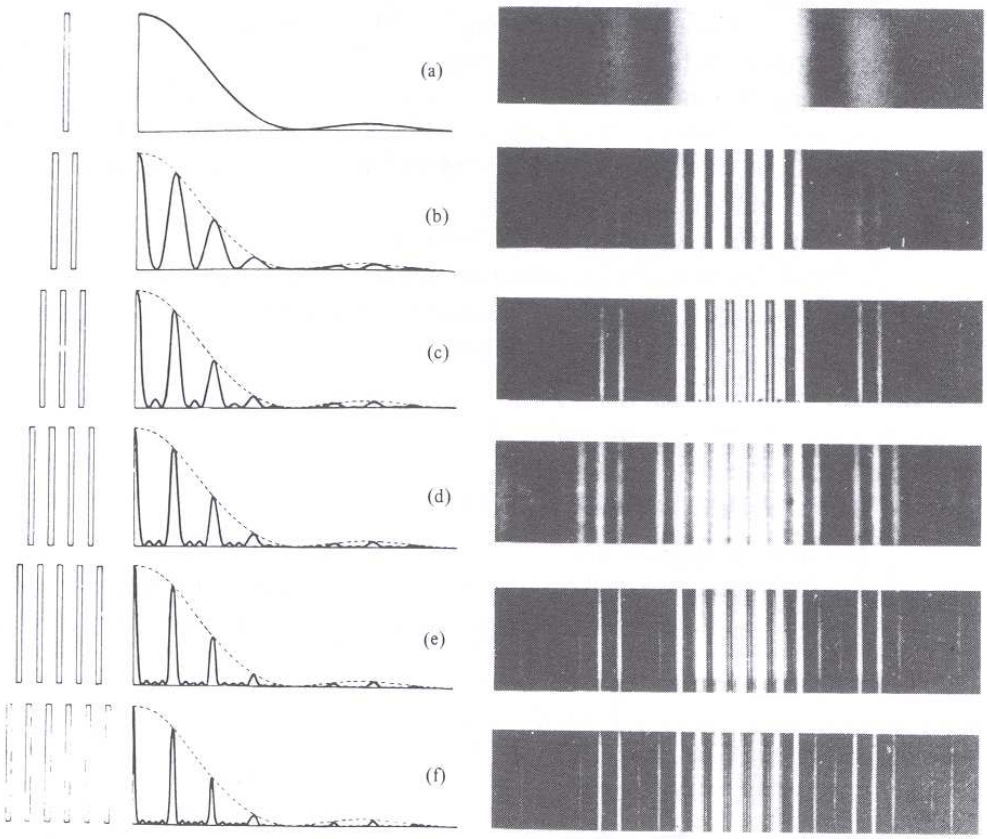


Figure 1-2 Fraunhofer diffraction form N-slit gratings

We won't detail here all the mathematical details as they can be found in many physics and crystallographic textbooks.[1, 2]

To show the concept described above an optical diffractometer can be set up.[3] An helium-neon laser is commonly employed. The diffraction pattern (Fourier transform) of a mask at O is formed at D, the focal plane of the objective L_1; and the image (double Fourier transform) of O is formed at I.

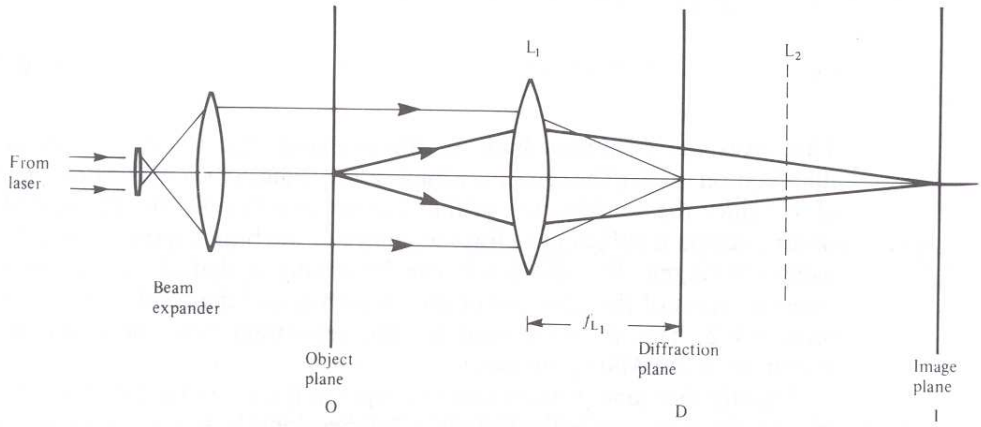
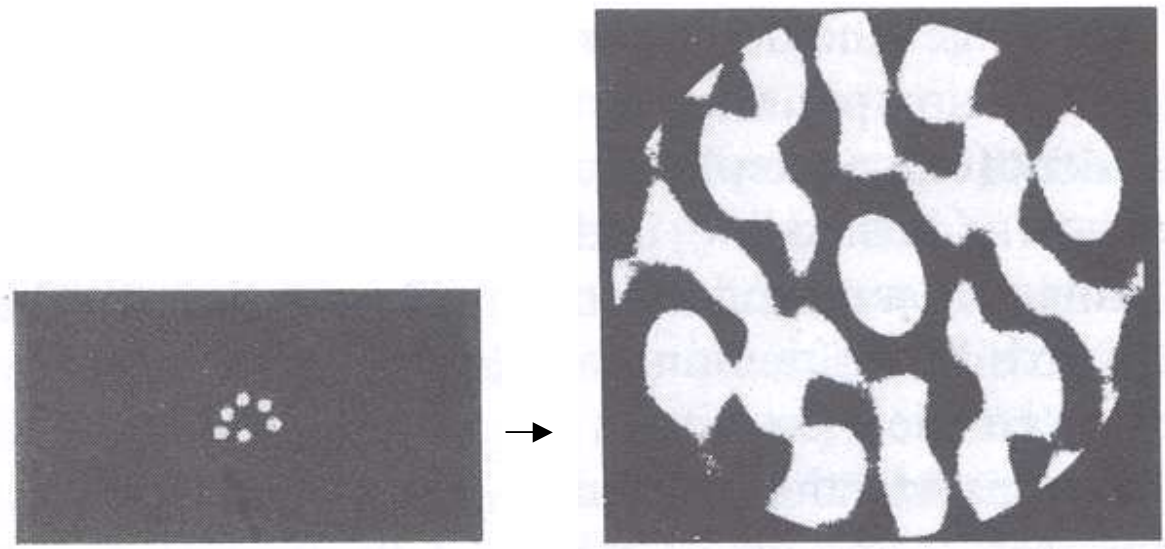


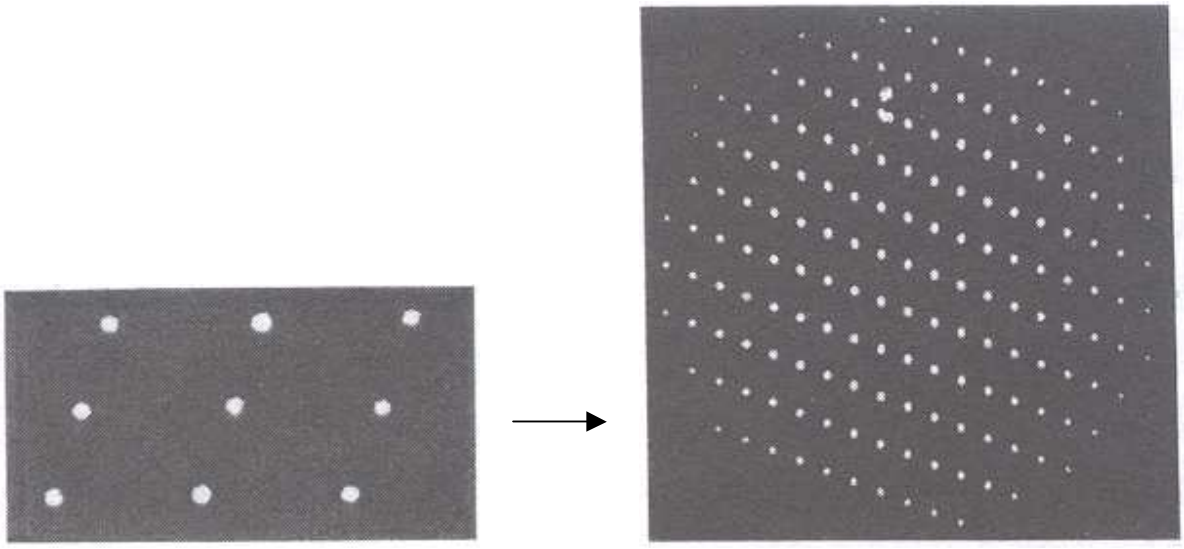
Figure 1-3 Basic arrangements for observing optical transforms

Let's begin using on object mask (a) consisting of a non-repeating irregular group of holes. This can be thought of a group of atoms in an isolated asymmetric unit and we can see its diffraction image (b):



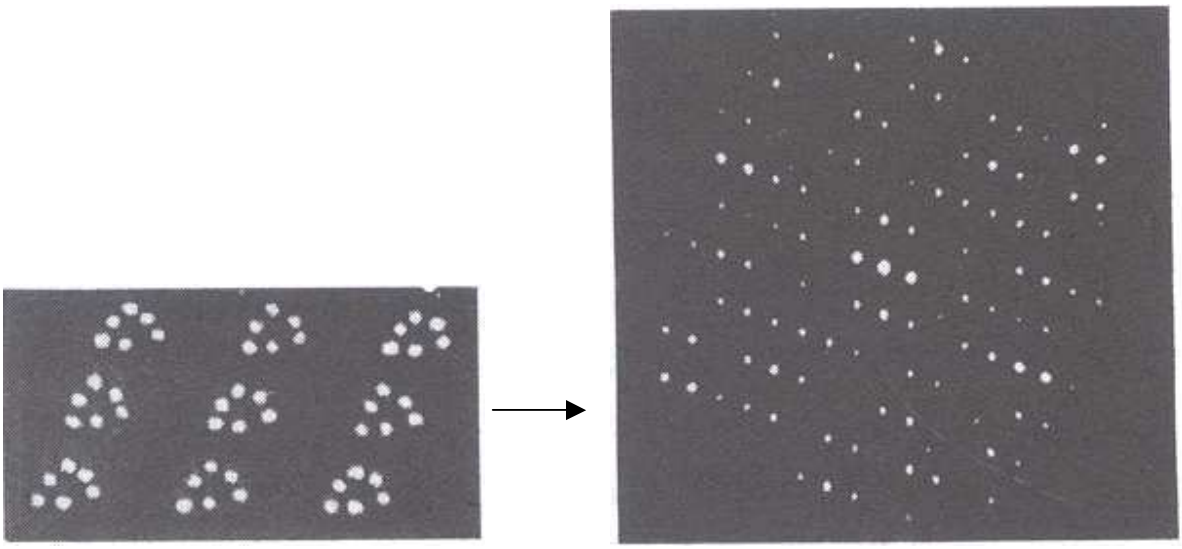
(a) (b)

If we take instead as object mask (c) where the holes are periodically repeated to form a bi-dimensional grating (in reality there are many more holes than the ones represented in the figure) we see that its diffraction image (d) is discrete exactly as in the case of interference of N-waves in the case of N very big:



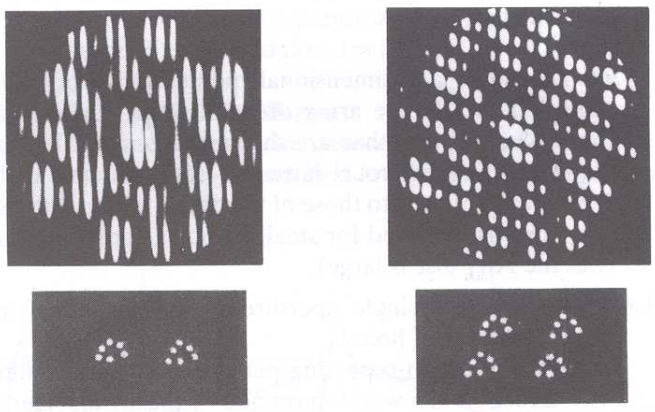
(c) (d)

If we take an object mask (e) that is the periodic repetition of the asymmetric unit with the same periodicity of the grating above we see that its diffraction (f) image is a sampling of the pattern (b) following (d), as we have seen in Figure 1-2:



(e) (f)

Other two examples:



The position of the diffraction maxima will be inversely proportional to the vectors a, b and c that define the crystal lattice. The direction of the maxima of interference is due to the periodicity present in the crystal. It is the lattice that determines the direction of the maxima and actually not even an atom needs to be located at a lattice point. In the simplest case of the interference of N sources along a line, the maxima will occurs at a position inversely proportional to the distance of the sources, in this case, as we consider a three dimensional case, there are three conditions of this kind to be fulfilled. Let's consider the following picture:

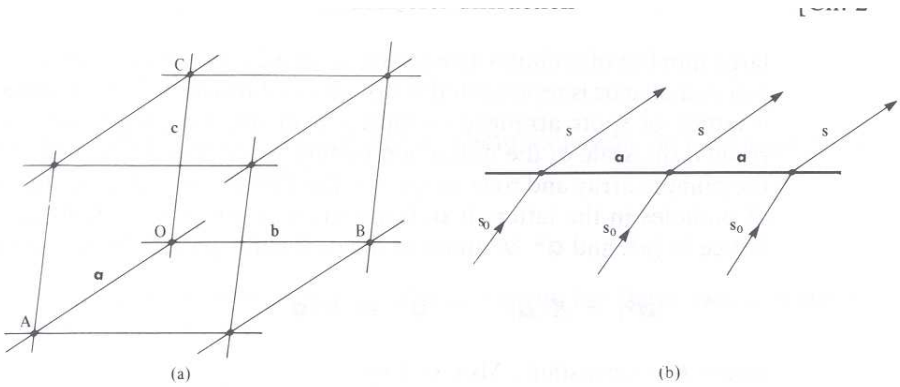


Figure 1-4 A crystal lattice as a 3-dimesnional diffraction gratings

and consider the **a** direction. Let us define $\mathbf{s_0}$ as the incident wave of wavelength λ and **s** the scattered one, each **s** will be in phase with its neighbour and we will observe a maximum if

$$\vec{a} \cdot \vec{s} - \vec{a} \cdot \vec{s}_0 = h\lambda \tag{1.1}$$

A similar reasoning must be done for the other two directions and at the end we have that in order to have a maximum these three conditions must be satisfied:

$$\begin{aligned} \vec{a} \cdot \vec{s} - \vec{a} \cdot \vec{s}_0 &= h\lambda \\ \vec{b} \cdot \vec{s} - \vec{b} \cdot \vec{s}_0 &= k\lambda \\ \vec{c} \cdot \vec{s} - \vec{c} \cdot \vec{s}_0 &= l\lambda \end{aligned} \tag{1.2}$$

eq (1.2) are called the Laue equations.

It can be shown that there is a complete equivalence between the Laue interpretation and the Bragg one.

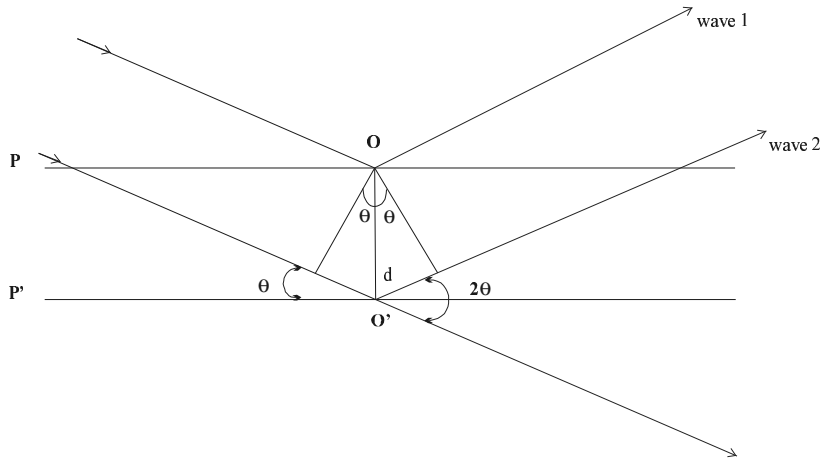


Figure 1-5 Bragg's conditions for diffraction

$$n\lambda = 2d_{hkl} \sin\theta \tag{1.3}$$

The scattered X-rays that cause a maximum of diffraction according the Laue equations also constitute a maximum in the optical sense, the reflection being at lattice plane defined by hkl values from the Laue equation.

The directions of the X-rays' diffraction maxima from a crystal are determined, for a given wavelength, by the dimension of the crystal gratings. By measuring those directions the lattice geometry can be determined.

To compute the intensity of the maxima we can derive the expression thinking in terms of Bragg reflection. As we already said we have to think of a crystal as a continuous distribution of scattering matter. Consider with the following figure the projection of the plane hkl=100,

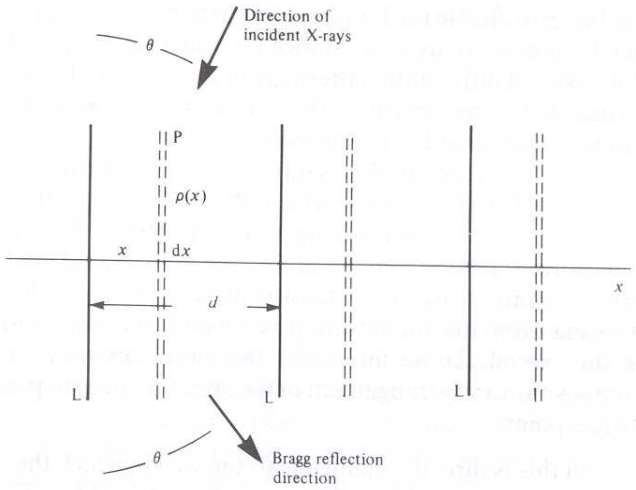


Figure 1-6 Bragg's reflection. Edge-on view of lattice plane

the first order reflection, with the lattice spacing d, corresponds to the lattice cell *a*. The scattering from the electron density that is in the plane is all in phase in the direction of the reflection as we noted above, but we must take into account the density between these planes. For the first order reflection there is a path difference of one wavelength (a phase difference of 2π) in the scattering of two successive planes. The contribution of the same reflection from the scattering plane P is therefore out of phase by $(x/d)\ 2\pi$, if $\rho(x)$is the

amplitude of scattering from a layer of thickness *dx* at *x*, the total amplitude of the reflection from all scattering matter is given by:

$$A_1(\theta) = \int_0^d \rho(x) e^{i(x/d)2\pi} dx \tag{1.4}$$

For the second order reflection, the hkl=200, there is a two wavelength path difference in the reflection from successive planes and the phase difference for plane P will be then *(2x/d) 2π*. In general for a reflection h00 the phase difference will be *(hx/d) 2π*. We can then write:

$$A_n(\theta) = \int_0^d \rho(x) e^{i(nx/d)2\pi} dx \tag{1.5}$$

and generalizing in three dimensions:

$$F(\vec{h}) = \int_V \rho(\vec{r}) e^{i2\pi(\vec{h}\cdot\vec{r})} d\vec{r} \tag{1.6}$$

with $\vec{r} = (x, y, z)$ and $\vec{h} = h, k, l$. Eq(1.6) is called the structure factor.

In analogy of the single-slit case we have found that the wave F(h) is the Fourier transform of the electron density and we can write as well that:

$$\rho(\vec{r}) = \int_{-\infty}^{\infty} F(\vec{h}) e^{-i2\pi(\vec{h}\cdot\vec{r})} d\vec{h} \tag{1.7}$$

the electron density is the Fourier transform of the structure factors..

Using in eq.(1.6) a continuum of electron density is not practical from a computational point of view. That is why the density integral over the cell is transformed into a summation over all the N atoms present in the unit cell. In this way we introduce the atomic scattering factor *f* to define the scattering power of each different atom. The X-ray scattering power will depend on the number of electrons in the particular atom and it decreases with increasing scattering angle and is higher for heavier atoms. For zero scattering angle the value of *f* equals the number of electrons. Introducing the scattering factor Eq.(1.6) will became:

$$F(h, k, l) = \sum_{j=1}^{N} f_j e^{i2\pi(hx_j + ky_j + lz_j)} \tag{1.8}$$

x, y and z are the atom fractional coordinates .

The scattering factor contains additional contributions from anomalous dispersion effects (essentially resonance absorption), which become substantial in the vicinity of the X-ray absorption edge of the scattering atom.

$$f = f^0 + f' + if'' \tag{1.9}$$

Another point is that due to the thermal motion the atoms oscillate around the positions of minimal energy and this will modify the electron density function and consequently their capacity to scatter. In order to take into account this effects a temperature factor is introduced

$$f_B = f \cdot e^{\left(-8\pi^2 U \sin^2 \theta / \lambda^2\right)} \tag{1.10}$$

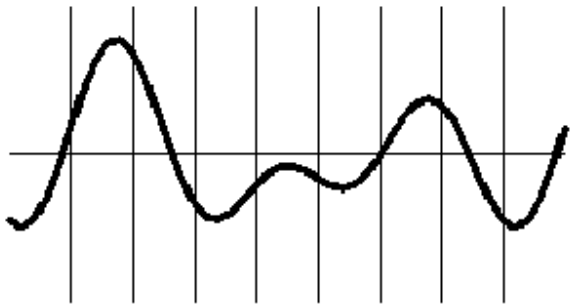
where U is the mean square shift of the atoms with respect to the position of equilibrium. Finally we can write the structure factor as:

$$F(h,k,l) = \sum_{j=1}^{N} f_j e^{i2\pi\left(hx_j + ky_j + lz_j - 8\pi^2 U_j \sin^2 \theta / \lambda^2\right)} \tag{1.11}$$

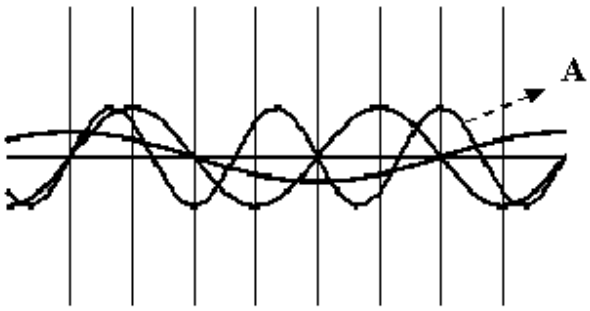
1.1.1 The phase problem

Eq (1.7) gives us the electron density, that is the information that we want to obtain after having performed the experiments. The refractive index of X-rays is very near to unity and for this reason they can't easily be focused by means of lenses. We can just compute the value of the density but now there is the main complication. The detectors can only count the number of photons that are diffracted and we can't have any information about the phase of the diffracted wave. We can collect only intensity values that are proportional to the square of the structure factors. Just to recall again in the case of ordinary light, it is the phase difference that creates the image, but in this case a lens is doing physically the integral of eq(1.7). As the image is crated due to this phase difference we can already intuitively think that is the phase term $e^{i2\pi(hx_j+ky_j+lz_j)}$ in eq (1.8) that plays the important role, the one that we can't measure. Because the electron density is seen as the superposition of an infinite

number of waves, the phase difference is really important in the summation if we have for example a function like this one

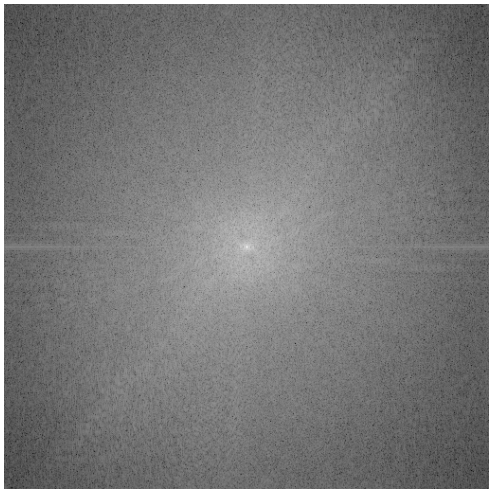


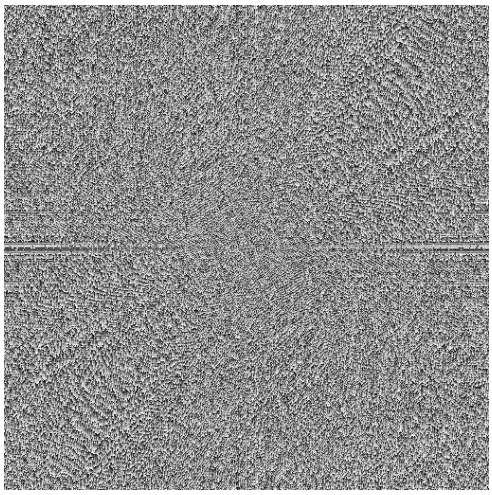
that can be obtained as the superposition of the following waves



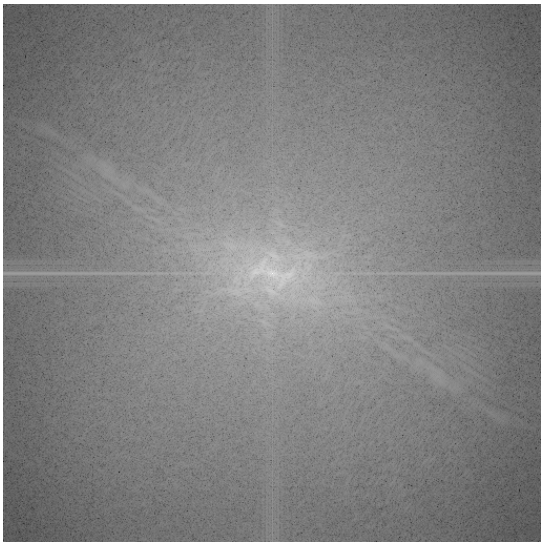
If we shift then for example the wave A on the right (we change its phase) we are not able anymore to obtain our starting function summing the four waves.
A more visual example is given with the following images[4]:

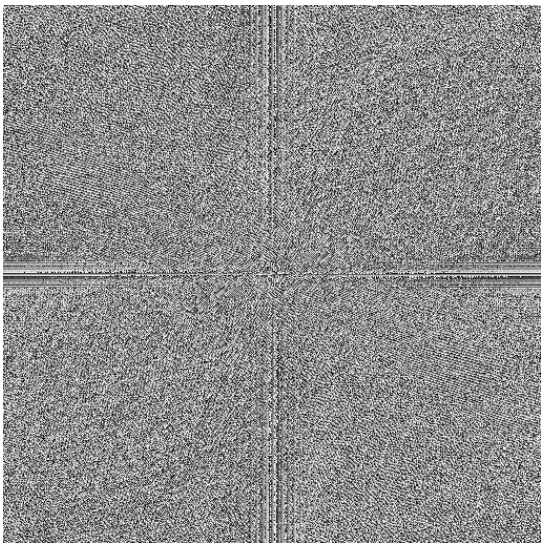
a cheetah picture


the magnitude transform of the cheetah picture


the phase transform of the cheetah picture

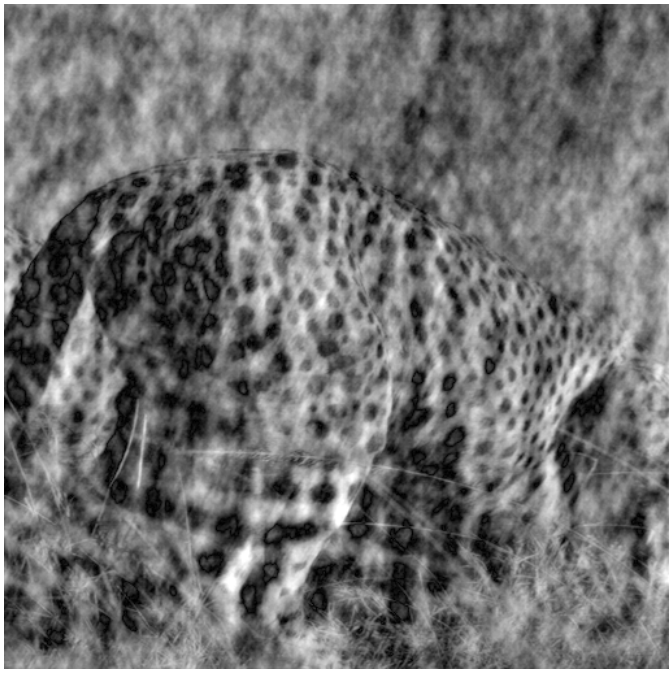
a zebra picture


the magnitude transform of the zebra picture


the phase transform of the zebra picture

Reconstruction with zebra phases and cheetah amplitudes:

Reconstruction with cheetah phases and zebra amplitudes:



The phases play the most important role.

Guessing the phases is not a good approach for a trial and error algorithm because if we think of a centrosymmetric structure where the phase can have only the value 0 or 180°, ignoring the anomalous scattering, and in only one dimension the number of different combinations is 2^n, where n is the number of reflections. The phase can have only those two values because when $\rho(x)$ is an even function its Fourier series presents only the *cos* terms because the ones with sin are equal to zero. When we have a hundred or a thousand reflections is impossible to guess the solution. Let us suppose to have 100 reflections, we will have 2^{100} possible combinations to search and if we can for example test 1000 combinations per second it will take almost $4*10^{19}$ years to solve the problem.
To overcome this problem, Direct Methods[1] have been invented that attempt to derive the structure factor phases by mathematical means from a single set of X-ray intensities. This method works for single crystal data where it is possible to have for each reflection its intensity value. From powder diffraction data this is not straightforward due to the overlap of the peaks and as it's difficult to obtain a good set of intensities, Directs Methods often fail with powder data. We have also often a too limited range in $\sin\theta/\lambda$ (d spacing to 1-2 Å) We'll introduce in the next paragraph the main peculiarity of a powder diffraction experiment.

1.2 Powder diffraction

A powder or polycrystalline sample ideally consists of a very large number of randomly oriented micro-crystals with dimensions ideally of approximately 1-10μm. The principles that govern the diffraction of single crystals also apply for powders; but, because of the random orientation of the crystallites, there are many planes *hkl* – each one with a different orientation coming from a different micro-crystal– that simultaneously fulfil Bragg's condition, thus producing a cone of diffraction with semi-vertical angle $2\theta_H$.
We have a collapse of the three dimensions of crystallographic information onto the single dimension of the powder diffraction pattern[5, 6].

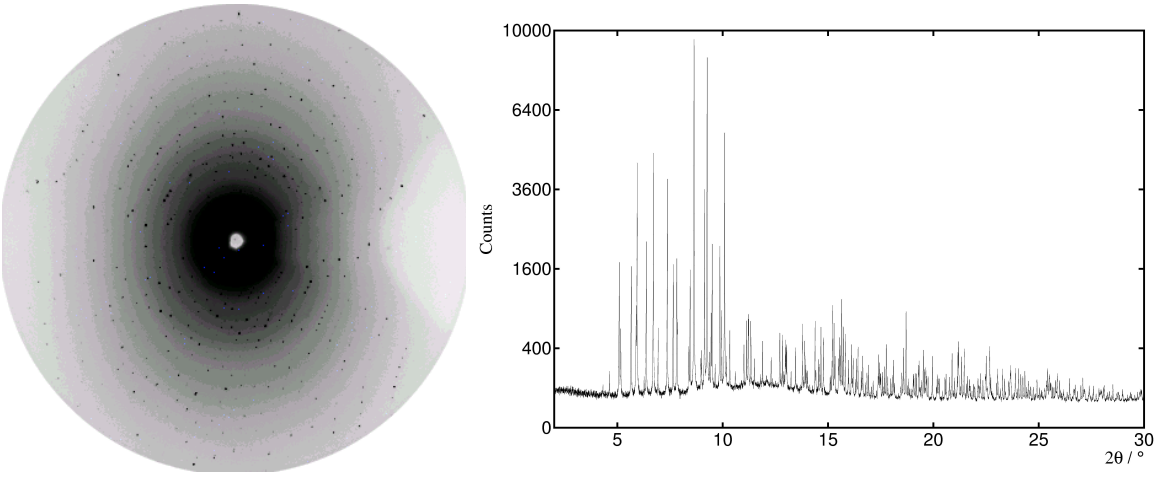


Figure 1-7 Comparison between a single crystal measurement (left) and a powder diffractogram (right)

This causes the problem of peak overlap. We can have exact overlap, for example in a cubic system the reflections 511 and 333 have the same d spacing and occur at the same 2θ angle, or inexact overlap as shown in the next picture when the d-spacing of the reflections are similar.

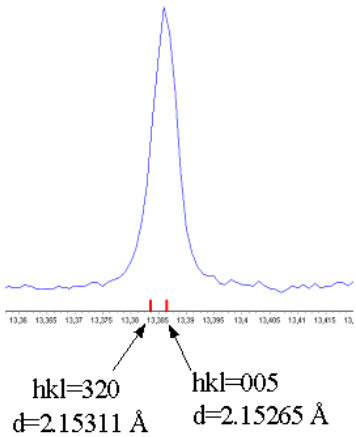


Figure 1-8 Peaks' overlap in powder diffraction

.

This is the cause of the two major difficulties comparing to an experiment with single crystal data. Due to the overlap, if we are dealing with an unknown structure it is difficult to obtain the crystallographic cell because it is not easy to understand how many reflections contribute to a single peak and what is their exact position. The other problem is that without a structural model it is not possible to extract the individual intensities, and this will be better explained with the help of a few formulae in section 1.2.5.

1.2.1 The Rietveld Method

In 1967 and 1969, H. M. Rietveld suggested a new approach to the analysis of powder diffraction data[7-9]. Instead of using the individual integrated intensities, the whole diffraction pattern was incorporated into the least squares refinement, minimising the difference between the observed pattern and the one calculated by modelling the instrumental, spectral and structural contributions to the pattern. This method, which was initially developed for neutron diffraction, was extended to conventional X-rays and synchrotron radiation.

If we have an initial model, we can calculate each point of our diffractogram, y_{ic} , in the following way:

$$y_{ic} = s\sum_{k} m_k L_k \mid F_k \mid^2 G(\Delta\theta_{ik}) + y_{ib} \qquad (1.12)$$

the sum over k is extended to all the overlapping peaks that contribute to the point i. In eq(1.12):

s is the scale factor

m_k is the multiplicity term (for eg. in cubic system the reflections 100, 010 and 001 have the same d-spacing and then they occur at the same angle and are related by symmetry)

L_k is the Lorenz-polarization term

F_k is the calculated structure factor

$G(\Delta\theta_{ik})$ is the profile function, y_{ib} is the background contribution

We could even insert more corrective terms for preferred orientation, crystal size effects and micro-strain. In our analysis we weren't interested about crystal size effects and micro-strain and we described the profile function from a completely analytical point of view. If severe preferred orientations are present in the sample it is almost impossible to solve a structure ab-initio.

The model is then refined minimising by a least squares process the residual:

$$S = \sum_i w_i \mid y_{io} - y_{ic} \mid^2 \qquad (1.13)$$

$(w_i)^{-1}$ is the sum of the standard deviations associated with the peak and the background

y_{io} is the observed i point

y_{ic} is the calculated intensity at the i-point with eq. (1.12)

As eq (1.12) is a non-linear function, the residual S of eq(1.13) will have many local minima. If the initial model X^0 is almost correct we can expand in a Taylor series eq.(1.12)

$$y_i(\mathrm{X}) \cong y_i(\mathrm{X}^o) + \sum_{j=1}^{m} \left(\frac{\partial y_i}{\partial x_j}\right)^0 \partial x_j + \frac{1}{2} \sum_{j,p=1}^{m} \left(\frac{\partial^2 y_i}{\partial x_j \partial x_p}\right)^0 \partial x_j \partial x_p + ... \qquad (1.14)$$

and we can omit in the expansion terms bigger than the first order

$$y_i(\mathrm{X}) \cong y_i(\mathrm{X}^o) + \sum_{j=1}^{m} \left(\frac{\partial y_i}{\partial x_j}\right)^0 \partial x_j + e_i \qquad (1.15)$$

in this way we can approximate as linear the function (1.12) and perform a linear least squares procedure. In this case we can't compute directly the values of our unknown parameters but instead we can compute the shift ∂x_j to apply to the starting values of the parameters of the initial model. This shift is computed until convergence is reached with the iterative procedure of eq 1-68 rewritten here in matrix notation

$$\Delta Y = A\,\Delta X + E \longrightarrow X = X^0 + \Delta X$$

In Figure 1-9 is schematically represented a one dimensional slice of the n-dimensional parameter space.

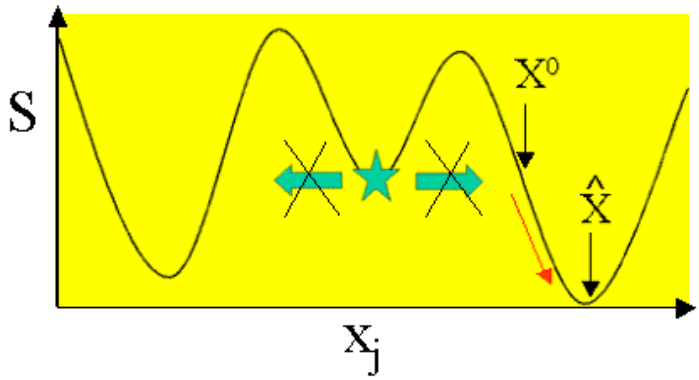


Figure 1-9 One dimensional slice of the n-dimensional parameter space

If our starting model is not right, for example we can start with a configuration near the local minimum represented with a star in the figure, there is no way that the least squares procedure brings the model to the true minimum but it can only converge to the local minimum.

If the model is unknown, in order to find a good starting model we can try to extract the integrated intensities from the powder diffractogram and use direct methods as in a single crystal experiment to solve the phase problem, or we can perform a real space search with Monte Carlo algorithms.

In any case if the structure is unknown we need to find as a first step the crystallographic unit cell.

1.2.2 Autoindexing programs

In a powder diffractogram we can only discern reflections with different d-spacing but we lose all the information about the direction of the diffracted beam that is present in a single crystal experiment. Finding the unit cell dimensions becomes more difficult and the following relation between the observed d-spacings and the reciprocal lattice vectors forms the basis for a large number of powder-indexing programs:

$$\frac{1}{d_{hkl}^2} = h^2a^{*2} + k^2b^{*2} + l^2c^{*2} + 2hka^*b^*\cos\gamma^* + 2klb^*c^*\cos\alpha^* + 2hla^*c^*\cos\beta^* \quad (1.16)$$

defining:

$$Q_{hkl} = \frac{10^4}{d_{hkl}^2} \quad (1.17)$$

we have:

$$Q_{hkl} = h^2a_{11} + k^2a_{22} + l^2a_{33} + hka_{12} + kla_{23} + hla_{13} \quad (1.18)$$

For each observed Q-value we need to find three crystallographic indices (*hkl*) , for which

$$Q_{obs} - \Delta < h^2a_{11} + k^2a_{22} + l^2a_{33} + hka_{12} + kla_{23} + hla_{13} < Q_{obs} + \Delta \quad (1.19)$$

where Δ is the Q_{obs} error.

The conditions of eq. (1.19) are necessary but not sufficient for a physically plausible index. In fact every powder diffractogram can be indexed with a huge cubic cell: with a cell like this it is always possible by chance that a calculated reflection matches an observed one, but we will have also a huge number of predicted reflections that are in fact not observed. To discard this kind of solution, figures of merit have been introduced. The de Wolff figure of merit[10] is based on the following expression:

$$M_{20} = \frac{Q_{20}}{2 \cdot \langle Q \rangle \cdot N_{20}} \quad (1.20)$$

where M_{20} is the number of calculated Q values up to Q_{20}, which is the Q value for the 20th observed indexed line; $\langle Q \rangle$ is the average discrepancy in Q for these 20 lines. The bigger is the value of M_{20} the more plausible is the solution because it means that there are fewer predicted lines that are not observed. Most of the time some predicted lines are not observed because of the systematic absences of the space group. High resolution is needed because decreasing the peaks' overlap we can have a better set of Q_{obs} and reduce <Q>. This is the main important step in all the procedure. If the data are of bad quality it is almost impossible to find a cell. Impurity lines cause also problems because they introduce into eq (1.19) wrong values; some of the programs can take this into account as well as the zero error.

Different programs are available that use different algorithms that can succeed in some cases or in others. All these programs have as output a list of different cells listed in decreasing order of figure of merit. Looking at the dimensions of the cells and at the compound density

it is possible to have a guess on which is the right cell and determine the space group. Only with a final Rietveld refinement can we be completely sure about our choice. A list of programs we used to index the powder diffractograms is the following:

- CRYSFIRE[11] that acts as an interface with the programs:

Dicvol91[12] (the reference is the newest version 04 not implemented yet in crysfire); Ito12[13]; Treor90[14]; Taup[15]; Kohl[16]; LOSHFZRF; Lzon612;FJZN[11]

-McMaille [17]

-Topas SVD indexing[18]

1.2.3 Extracting integrated intensities

Once the unit cell has been found one can think to extract the integrated intensities from the powder diffraction pattern and give them to a Direct Method program as with single crystal data. Two principal techniques have been developed: the iterative LeBail method and the Pawley method.

1.2.3.1 The iterative Le Bail algorithm

The Le Bail method[19] is a recursive version of the original Rietveld approach. In the Rietveld method if a structural model is known it is possible to estimate the structure factor magnitudes. For overlapping peaks in fact, the contribution for a given reflection is weighted by the calculated peak contribution for that reflection divided by the sum of the calculated peak values for each contributing reflection. The calculated peak contribution is coming from the model, so if we have a model it is possible to extract the integrated intensities. If the structure is not known we can start with the assumption that all the integrated intensities are equal and then recursively solve the formulae that give us the integrated intensity in the Rietveld code:

$$A_m^{r+1}(obs) = \sum_i \frac{A_m^r(obs) \times q_m(i)}{(\sum_{n=1}^{N} A_n^r(obs) \times q_n(i))}(obs(i) - back(i)) \qquad (1.21)$$

where

$$q_m(i) = c_m(i)G(\Delta\theta_{im}) \tag{1.22}$$

$c_m(i)$ contains Lorenz-polarization, absorption and extinction terms and $G(\Delta\theta_{im})$ is the peak shape. In the denominator the sum over the N terms is extended over the N peaks that contribute to that reflection. The method is stable but problems can occur for weak peaks for which the background can be overestimated and the procedure can't reach convergence. In many cases it is better to determine the background correctly and leave it fixed.

1.2.3.2 The Pawley Method

The Pawley method[20] is a least-square procedure where the variables are the peak areas themselves. Every point in the powder diffractogram is calculated as:

$$y_{ic} = back(i) + \sum_k A_k q_k(i) \tag{1.23}$$

where the symbols are the same as in eq.(1.21).

The integrated intensities are evaluated minimizing the summed weighted difference between the observed and the calculated diffractogram.

$$S = \sum_i w_i \mid y_{io} - y_{ic} \mid^2 \tag{1.24}$$

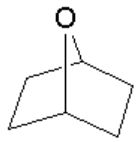
This is a simple case as eq (1.23) this time is linear and the solution may be achieved by a single matrix inversion if the peak shape function is kept fixed. Things are not so straightforward because the matrix is degenerate in the case of exact overlap and ill-conditioned for close peak separation. To overcome this problem reflections that are closer than a particular criterion such as a quarter of a FWHM can be grouped together. Close peak separation leads to the appearance of negative intensities. To avoid this, Waser-type constraints [21] can be used to minimize the differences between neighbouring integrated intensities.

1.2.4 Real space search with Monte Carlo algorithms

When we try to solve a structure, if we are able to determine the unit cell and the space group, we have already a lot of information and most of the time we know which atoms or which molecules form the crystal. Looking at the volume of the cell, it is possible to determine how many atoms or molecules form the unit cell, and by looking at the space group it is possible to see how many are independent and form the asymmetric unit. Once this is determined it is possible to simulate trial configurations of our crystalline structure in order to maximize the agreement between the calculated and the observed diffraction pattern. No attempt is made to extract individual intensities and treat them as in single crystal data.

We can think of solving the problem using a *grid search* algorithm in which all the atoms are moved systematically around over the x,y,z, space of the asymmetric unit on a number of discrete grid points in such a way that all the possible configurations are explored. If we divide our asymmetric unit into a 10*10*10 grid we have 10^3 possible locations for each atom and if we have a simple structure with 7 atoms we can have, ignoring the possibility of atomic overlap, 10^{21} atomics configurations. If we can evaluate 1000 possible structure every second it will take almost $3*10^{10}$ years to finish the computation. We cannot think to solve the problem in this way.

The first problem is the description of the crystal structure. We simply described our structure as formed by independent atoms but in realty the atoms are bonded to each other. If for example we take the 7-oxabicyclo[2.2.1]heptane that is formed by 7 atoms and we describe this molecule as a rigid body



we only have 6 parameters that can vary, 3 for the positions and 3 for the orientations. We only have now 10^6 possible configurations that will be explored in almost 17 minutes.

If the structure is not completely rigid and there are additional degrees of freedom due to the torsion angles, then the problem becomes again more complicated and not soluble with a grid search algorithm. For example for a rigid body with 10 torsion angles we must explore 10^{16} different configurations and it will take 32000 years to explore them all. The grid search is not a good algorithm to solve this problem.

1.2.4.1 Simulated annealing

As seen in paragraph 1.2.1, with the Rietveld method the structural parameters can be adjusted in the direction of decreasing residual S, eq(1.13). If our model is wrong we can only locate the local minima. We need a method that is more efficient and allows us to escape from a local minimum and that is able to move in directions where the residual S is increasing. The simulated annealing[22, 23] algorithm has been introduced to try to solve this problem.

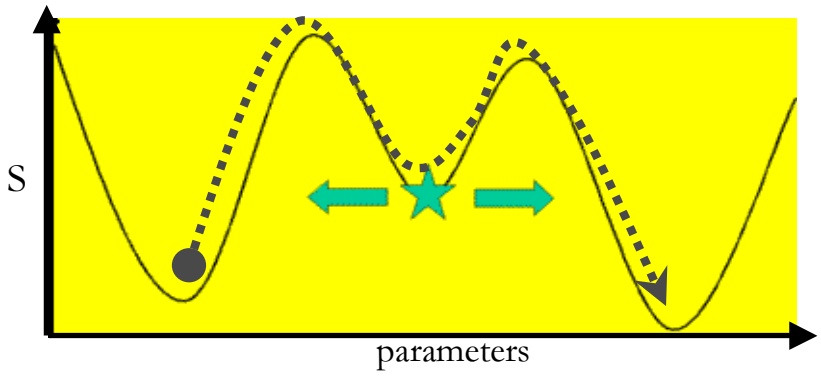
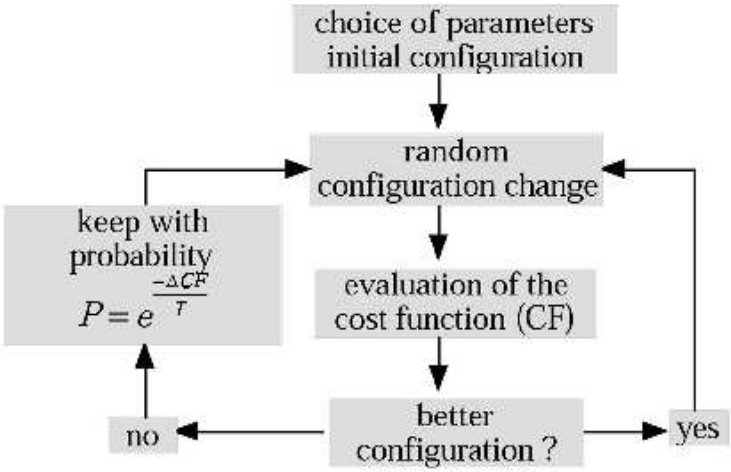


Figure 1-10 One dimensional slice of the n-dimensional parameter space

To see if the agreement between the observed and the calculated data is decreasing or increasing we need to define a cost function (CF). Different algorithms use different CF but the common workflow of the simulated annealing algorithm is the following:



Starting from an initial random configuration the free parameters are randomly varied. The CF is then computed, if its value its lower than before the randomization then this configuration is kept and another randomization is done and so on. If the CF is bigger than the value before the randomization then this configuration is kept with a probability $P = \exp[(CF_2-CF_1)/T]$, where CF_1 is the cost function of the initial configuration and CF_2 is the cost function after the randomization. This means that the system is allowed to perform moves uphill and has the possibility to escape from local minima. The probability to accept a worse configuration is proportional to the parameter T that is called temperature; the higher the temperature, the higher is the probability to keep the configuration. That means that at high temperature the system is explored thoroughly. To make the process converge it is necessary to decrease the temperature, but if the temperature is decreased too quickly there is the risk to trap the system in a local minimum. Another important point concerns the randomization of the parameters. If the free parameters are randomly varied a lot in each randomisation, there is a very small probability to find a better configuration and so it is preferred to vary the parameters within a small range. A so-called Markov chain is performed where each parameter is varied a little from one randomization to another but after having performed millions of randomizations we loose any link with the initial configuration.

The fundamental parameters in the algorithm are:

- the definition of the CF that defines the hypersurface that has to be explored;
- the variation of the temperature during the optimisation;
- the amplitude of the displacements;
- the description of the molecule, where using chemical information it is possible to describe the structure in a better way than simply as isolated independent atoms.

We will see in detail how these parameters are defined in the program DASH[24] that is the one that we used to perform the simulated annealing optimization.

In DASH a previous Pawley fit is done in order to extract the integrated intensity without the structural model. In DASH those intensities are not treated as in a single crystal data. From a Pawley fit we obtain the covariance matrix that gives us the correlation between overlapping peaks and this information is used in the algorithm. The CF is defined in the following manner:

$$CF = \sum_h \sum_k [(I_h - c \mid F_h \mid^2)(V^{-1})_{hk}(I_k - c \mid F_k \mid^2) \qquad (1.25)$$

where $I_{h,k}$ is the extracted intensity from a Pawley refinement of the diffraction pattern, V_{hk} is the covariance matrix from the Pawley refinement, c is the scale factor and $F_{h,k}$ is the calculated structure factor from the current trial structure. In this way we take into account the peaks' overlap and we increase the computational speed because we avoid computing every point of the diffractogram as in the Rietveld method. If the search space is complex (i.e. contains many local minima), the starting temperature must be high so that the algorithm will make many uphill moves in the early stages. This will ensure that the search space is explored thoroughly. DASH selects the starting temperature by performing a brief simulated annealing run at high temperature and monitoring the variance of the CF value as random moves are made in parameter space. Based on the results, the starting temperature is set to a value that will allow the algorithm to escape deep local minima at the outset of the search (i.e. the higher the variance in CF, the higher the starting temperature). DASH typically uses a fixed, conservative cooling rate of 0.02 K (i.e. the rate at which the temperature is reduced) for the annealing process. The temperature reduction is applied at the end of each cycle of

annealing. The cooling rate is not constant as the annealing proceeds. If DASH detects large fluctuations in the CF (implying that the algorithm is in an interesting region of parameter space) it automatically reduces the cooling rate to ensure a thorough search.

$$T_{next} = \frac{T_{current}}{1 + [CT_{current} / 3(\langle CF^2 - \langle CF^2 \rangle \rangle^2)^{1/2}]} \tag{1.26}$$

where C is the cooling rate. The slower the cooling rate, the more extended is the search of parameter space and the greater the chances of finding the global minimum. However, a slow cool takes longer.

The parameters are adjusted individually and sequentially. The first parameter is adjusted once followed by the next and so on. The parameters are sampled from an exponential distribution $(\Delta x/\Delta x_0)^2 \exp\text{-}(\Delta x/\Delta x_0)$ where $\Delta x = x \pm \Delta x_0$, x is the current value and Δx_0 is a characteristic distance that is calculated to ensure an accept:reject ratio of around 50:50.

The molecule is described with a Z-matrix approach. Rather than using individual x, y, z coordinates to express the atomic positions, the Z-matrix is parameterised in terms of bond lengths, bond angles and torsion angles between atoms in the molecules. If everything is fixed the molecule is represented as a rigid body and there are 7 free parameters, 3 for the translation and 4 for the rotations as quaternions are used to describe the rotations. Is it possible to make the molecules more flexible by letting vary the values of torsion angle, bonds angle and bonds lengths and it is also possible to set limits to those values. The higher the number of free parameters the heavier becomes the computational cost. Describing the molecule with a Z-matrix creates a problem due to the torsion angle variation. If we vary a torsion angle at the beginning of the molecule within a small value we can create a big displacement on the atomic position at the end of the molecule (Figure 1-11) and this can cause problems because the probability of obtaining a better configuration in the case of a big displacement is low.

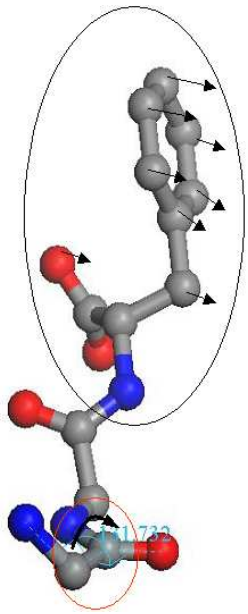


Figure 1-11 Representation of the molecule as a rigid body

The other problem is, as we already said, if the temperature is decreased too quickly then there is the risk to trap the system in a local minimum. To avoid this, multiple optimisations are done and if the process keeps converging to the same minimum after multiple simulated annealing runs, we can have increasing confidence in the solution.

1.2.4.2 Parallel tempering

To avoid the SA algorithm becoming trapped in a local minimum due to a too quick temperature decrease another algorithm called parallel tempering (PT)[25] has been introduced. In PT instead of using a single chain of configurations with decreasing temperature, a small number of optimisations are made, each with a different temperature. Periodically (eg. in FOX[26] - the program that we have used - , after performing ten trials in each parallel optimisation), the algorithm tests an exchange of configurations between parallel optimisations, with the same acceptation rules as when evaluating successive trials.

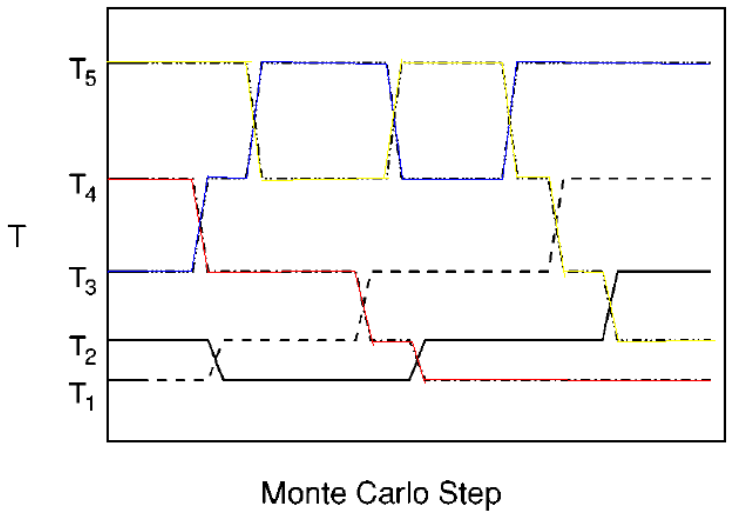


Figure 1-12 Evolution of configurations through parallel optimisation

There is no need to decide how many trials must be performed at each temperature and at all times there is the possibility to escape form a local minimum. In FOX thirty parallel optimisations are performed and no attempt to extract the intensities is done. There are available different CFs. In powder diffraction the most used CF is the weighted profile *R* factor:

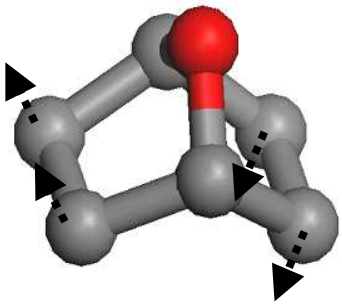
$$R_{wp} = \left[\frac{\sum_i w_i (y_{io} - y_{ic})^2}{\sum_i w_i (y_{io})^2} \right] \tag{1.27}$$

In the case of badly defined reflection profiles this CF doesn't work well and for this reasons in FOX the integrated R_{wp} is also available as CF. This is calculated by summing the total intensity scattered around each peak position in the range $2\theta_0$-f*FWHM to $2\theta_0$+f*FWHM, with f=1 for a Gaussian shape and f=2 for a Lorenzian. This range allows the inclusion of more than 90% of the diffracted intensity. If two reflections overlap the integration segments are divided between the two reflections and so no assumption is made about the actual values of the Bragg intensities. It is also possible to add an anti-bump CF that adds a penalty when two atoms are closer than a minimum distance. The anti-bump helps the algorithm to avoid improbable configurations but decreases the computational speed.

Also although in PT there is no need to define the cooling rate, it is important to find the range of temperature between the different optimisations. The high and the low temperatures

must be defined and also the amplitude of the random displacements in each optimisation. There are different possibilities that the user can chose and the following is the one that we used. Between the 30 different optimisations an exponential decrease of displacement amplitude values was chosen, from 1 Å down to 0.01 Å for translation, and from 36° to 0.36° for angles. The temperature values are then adjusted so that the same acceptance rate between 10% and 30% is reached.

Instead of describing the crystal structure with a Z-matrix, in FOX is used a "Molecule" representation. A molecule is defined by a list of atoms with their xyz coordinates (in an orthonormal reference frame internal to the molecule), with the geometry of the Molecule set by a list of restraints: bond lengths, bond angles and dihedral angles. The orientation of the molecule is defined internally by a quaternion, so its parameters are not directly accessible. Fox will allow atoms to move slightly away from the 'ideal' positions defined by the length and angular restraints and in this way it is possible to overcome the problems of the torsion angle in the Z-matrix previously described. FOX performs intelligent moves in order that the shape of the molecules is always kept.



1.2.4.3 TOPAS and the constrained conjugate gradient routine

The Conjugate Gradient (CG) method[27-29] is typically associated with solving the linear system $\mathbf{A}\,\mathbf{x} = \mathbf{Y}$ for large sparse symmetric positive definite $\mathbf{A}$ matrices. For sparse matrices CG is faster than LU-decomposition[30] and additional memory usage is minimized. CG is a successive approximation (ie. iterative) technique that minimizes on $(1/2)\,\mathbf{x}^T\,\mathbf{A}\,\mathbf{x} - \mathbf{Y}^T\mathbf{x}$. The CCG routine developed for TOPAS[31] modifies the CG routine to include min/max constraints. This has significant advantages over simply limiting parameter values after solving the system with LU-decomposition. Also of significance is the removal of parameters that are slow changing from the CCG loop. This significantly increases the speed of solution for dense or sparse systems such that CCG is 80 to 90 times faster than LU-decomposition for 1000 variables. We will not explain here the details of the CG as it is too long and a very clear report of 64 pages on this subject has been written by Jonathan Richard Shewchuk[32].
We will explain instead the different approach used in TOPAS for structure solution. This is not simulated annealing, genetic algorithms or direct methods. As previously seen with simulated annealing no information is used on the curvature of the cost function. A structure is simulated, the cost function is computed, and then the configuration is kept or not as explained in the previous paragraph. Simulated annealing will have difficulty in an extreme case like in figure 1-13

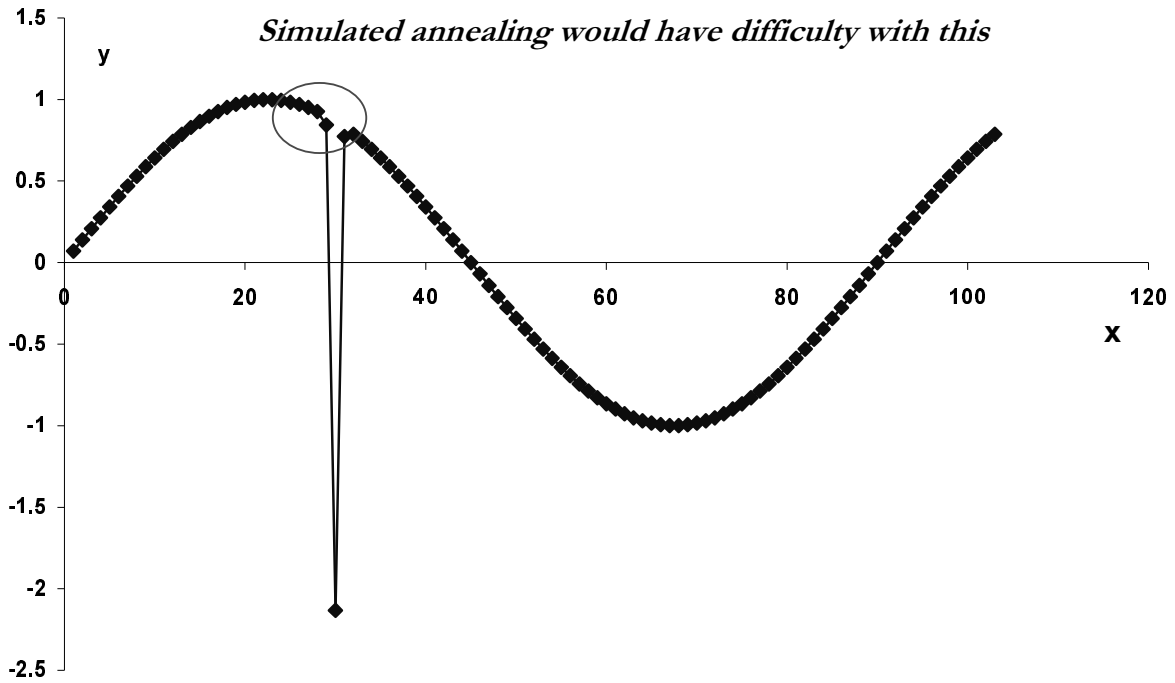


Figure 1-13 Kind of function where simulated annealing will have difficulties

Survival of the fittest loses significance and in the case of a minimum with a gradient that is very steep then a Monte-Carlo or pseudo-random search is as good as any.

In TOPAS after each randomisation the structure is refined and then every time information about the curvature is used; this is possible due to the CCG routine, much faster and stable than the LU-decomposition. In the example of Figure 1-13, starting the refinement from an initial configuration in the red circle, almost ten different starting points, would have been able to push the configuration straight to the true minimum.

In TOPAS there are different ways in performing the randomisation. Let us suppose we have a free variable x:

x @ 0.123 val_on_continue = Val + T Rand(-.1, .1)

the command val_on_continue = Val + T Rand(-.1, .1) means that after the first refinement, the value of x will be the actual value plus a random values between –0.1 and 0.1. After this randomisation another refinement is done.

Typically something like the following is used:

```
temperature 1 ' line 1
temperature 1 ' line 2
temperature 1 ' line 3
temperature 2 ' line 4
   use_best_values
   save_values_as_best_after_randomization
   move_to_the_next_temperature_regardless_of_the_change_in_rwp
```

Thus if the χ^2 did not reduce in the first three temperatures (lines 1 to 3) then at line 4 the best values found in lines 1 to 3 is reset. Randomization is performed with a greater temperature at line 4 in order to break out of the current parameter configuration. Line 4 also saves the new parameter configuration as the current best values. Thus a greater temperature at line 4 minimizes the chances of the process revisiting previous parameter configurations. A lot of these cycles are launched where the starting value is improving until the solution is found.

It is often difficult to formulate optimum val_on_continue functions by hand. This is especially true in structure solution using rigid bodies where optimum randomization of the rigid body parameters is difficult to ascertain. Randomize_on_errors is a means of automatically randomizing parameters based on the approximate errors in the parameters. When using this kind of randomisation, if the current parameter configuration is similar to a previous configuration, then the current refinement cycle is terminated and the temperature is moved on to the next one; this improves a lot the speed of the computation. Macros are available or can be created by the user in TOPAS.

Within the Auto_T(t) macros are calls to the macro TE_0(t); this macro contains the following:

```
temperature = (t);
    move_to_the_next_temperature_regardless_of_the_change_in_rwp
     on_best_rewind
  temperature 1
  temperature 1
  temperature 1
  temperature 1
  temperature 1
  temperature 1
```

t is typically set greater than 1. Thus a high temperature is initiated followed by a series of annealing temperatures. TE_0 is called six times within Auto_T and thus 6 little annealing runs are performed starting at different configurations due to 't' being larger than 1.

The command on_best_rewind is a little complicated but in effect it rewinds the temperature within the six annealing runs moving across the 't' temperatures as well.

After the 6 annealing runs comes:

```
temperature 1 use_best_values
```

This sets the current parameter configuration to the best values found within the 6 annealing runs. The subsequent temperatures of 1 as in:

```
load temperature { 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
                   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 }
```

simply anneals around the best values found in the 6 annealing runs.

The last temperature in Auto_T of:

```
temperature = 2 (t);
save_values_as_best_after_randomization
move_to_the_next_temperature_regardless_of_the_change_in_rwp
```

increases the temperature in order to break out of the current configuration.

It can be seen that the randomisation follows a precise scheme, ie. It is not completely random, and the parameters are adjusted, saved and then randomised again as the simulation is going on.

Performing the refinement after each randomisation is the longest process but what makes it competitive compared with simulated annealing is that if a configuration is near the true minimum, the algorithm is able to find the solution, in a few steps. Instead if the temperature in the simulated annealing is still high, the algorithm will continue to explore the N-dimensional space since it is not able to see whether the configuration is near the solution, as it is not sensitive to the curvature of the space but it will only be trapped in the true minimum at the end of the simulation when the temperature will be lower.

1.2.5 Fourier map

When we perform an experiment we are only able to collect a finite number of reflections. We can't use eq. (1.7) to define the electron density, but we should write instead

$$\rho(x,y,z) \cong \frac{1}{V}\sum_{hkl} F(hkl)e^{-2\pi i(hx+ky+lz)} \tag{1.28}$$

where the sum is extended over all the h, k, l observed reflections.

The density computed in this way is only an approximation because the sum should be infinite. This truncation causes the appearance of ripples and bubbles of density and decreases the resolution of the map.

The other very important point is that the observed map has a lot of information coming form the structural model that is not observed but obtained with simulations or with Direct Methods procedures. When we define the observed electron density as:

$$\rho_{obs}(x,y,z) = \sum_{hkl} F_{obs}(hkl)e^{-2\pi i(hx+ky+lz)} \tag{1.29}$$

we introduce the observed structure factor that is defined as:

$$F_{obs}(h,k,l) =\mid F_{obs}(h,k,l) \mid e^{i\varphi_{calc}(h,k,l)} \tag{1.30}$$

Because it is a complex number, it's defined by its modulus and its phases, but as we already said is not possible from the experiment to obtain the phases and so we use the calculated ones coming from the model

$$tg\varphi_{calc} = \frac{imaginary_part_F_{calc}(h,k,l)}{real_part_F_{calc}(h,k,l)} \tag{1.31}$$

the modulus instead is proportional to the intensity square root:

$$\mid F_{obs}(h,k,l) \mid = \sqrt{\frac{I_{hkl_obs}}{sm_k L_k}} \tag{1.32}$$

From a powder things are more difficult then from a single crystal experiment because also the intensities are model dependent due to the peaks' overlap. From a powder diffractogram we can extract the observed integrated intensity that is computed for an isolated peak as:

$$I_{hkl_obs} = \sum_i (y_{io} - y_{ib}) \tag{1.33}$$

where the sum is extended over the i-points that form the peak.

But in the case of overlapping peaks the net diffraction intensity is proportioned at a step *i* in the ratio of the contributions of overlapping reflections:

$$I_{hkl_obs} = \sum_i (y_{io} - y_{ib}) \frac{I_{hkl_calc}}{\sum_k I_{hkl_calc}} \tag{1.34}$$

The sum over k is extended to all the peaks that overlap and contribute to the intensity in the *i* point.

The extracted intensities are then model dependent because:

$$I_{hkl_calc} = \sum_i (y_{ic} - y_{ib}) = s \sum_i m_{hkl} L_{hkl} \mid F_{calc}(h,k,l) \mid^2 G(\Delta\theta_{ihkl}) \tag{1.35}$$

and

$$F_{calc}(h,k,l) = \sum_{j=1}^{N} f_j e^{i2\pi(hx_j + ky_j + lz_j)} \tag{1.36}$$

where x_j,y_j,z_j, are the fractional coordinates of the atoms in the model.

Using a difference Fourier map to complete a partial model from powder is then more difficult than from single crystal data because the maps that we can obtain are more model biased. This is why it is better to perform simulations and why often when a simulated model is wrong but it fits quite well the data it is difficult to improve it doing Fourier synthesis.

One thing can be improved; we can increase the resolution of the map avoiding the truncation effect of the Fourier series. This is what the Maximum Entropy Method is doing from which the density is computed in a completely different way without introducing the Fourier series.

1.2.6 Maximum Entropy Method

The standard way of refining a crystal structure from X-ray diffraction data is to approximate the electron density $\rho(\mathbf{r})$ in the unit cell of the crystal by a model electron density, which is parameterized by a relatively small number of parameters. Values of these parameters are then determined that correspond to the best fit to the measured dataset. This is possible thanks to the small number of parameters compared to the number of data. Usually some hundreds or thousands of reflections are available for the determination a few dozen to hundreds of parameters. Usually only one optimal solution exists. Another case occurs, if $\rho(\mathbf{r})$ should be described by its values in an arbitrarily dense grid in the unit cell.

The number of values to be determined can now be much higher than the number of the individual data. In such a case an infinite number of equally good solutions exists, and some additional criterion is needed to distinguish the best solution. Treatment of this problem leads to the Maximum Entropy Method (MEM).[33-38]

Let us suppose to divide our unit cell into a huge number of boxes with the same size. After this we take our electron density and we divide it into small equal pieces. We start to throw these pieces of density randomly and without any criterion into the boxes. There is no underlying bias in the procedure. When all the pieces have been thrown into the boxes we can see how many pieces are found in each box and this gives us a possible assignment for the value of the density in the ith box. Such density can then be accepted if it is consistent with the observed data otherwise it will be rejected. After more than one process of this kind has been done, we can empty again our boxes and start again from the beginning. After many such trials, some distribution $F(\{ni\})$ will be found to come up more often than others; the

one that most frequently occurs and is consistent with the observed data will be a sensible choice for the density. The process is completely random and we don't introduce any bias in the procedure. We can find in this way a density that is consistent with the observed data without introducing any artefacts. We will show in the next lines that this distribution corresponds to the greatest value of the entropy function, explaining again all the process in mathematical terms.

We need to introduce before a brief digression on permutations and combinations.

Let us consider in how many ways it is possible to arrange n different objects in a straight line: there are n choices for the first item, leaving n-1 possibilities for the second; this means that two of the n objects can be placed in n*(n-1) ways. For example if we have the numbers 1,2,3,4 we can place number 1 in the following places:

1 _ _ _ or _ 1 _ _ or _ _ 1 _ or _ _ _ 1

and number 2 can go in any of the 3 free places in total we have 4*3=12 different possibilities.

Continuing in the general case of n objects, there are n-2 choices for the third item and so on; in general we have that the total number of permutation is obtained by the product:

$$n \times (n-1) \times (n-2) \times (n-3) \times \ldots \times 3 \times 2 \times 1 = n! \qquad (1.37)$$

For example in the case of numbers 1,2,3,4 we have 4!=24 combinations:

1,2,3,4	1,2,4,3	1,3,2,4	1,3,4,2	1,4,2,3	1,4,3,2
2,1,3,4	2,1,4,3	2,3,4,1	2,3,1,4	2,4,1,3	2,4,3,1
3,1,2,4	3,1,4,2	3,2,1,4	3,2,4,1	3,4,1,2	3,4,2,1
4,1,2,3	4,1,3,2	4,2,3,1	4,2,1,3	4,3,1,2	4,3,2,1

A closely related question is in how many ways we can pick *m* objects from *n* different ones. In fact we have already answered this before; we just stop the product of eq.(1.37) when m items have been chosen.

$$n \times (n-1) \times (n-2) \times (n-3) \times \ldots \times (n-m+2) \times (n-m+1) = \frac{n!}{(n-m)!} = {}^{n}P_{m} \qquad (1.38)$$

We have seen before in the case of 4 numbers, than picking just two of them gives 4*(4-1)=12 possibilities.

We can pick:

1,2 1,3 1,4 2,1 2,3 2,4 3,1 3,2 3,4 4,1 4,2 4,3

If we are not interested in the order in which the m objects are picked, (i.e. the couple 1,2 and 2,1 are considered the same combination), this can be taken into account by dividing the result of eq.(1.38) by the number of ways in which m items can be permuted (i.e. as seen before, m!), this resulting number is denoted as:

$$^{n}C_{m} = \frac{n!}{m!(n-m)!} \tag{1.39}$$

$$^{4}C_{2} = \frac{4!}{2!(4-2)!} = \frac{24}{4} = 6$$

The sequence in fact now is:

1,2 1,3 1,4 2,3 2,4 3,1

Let us suppose now for example that we have two boxes (M=2), one is red the other is blue, and three pieces (N=3) that are randomly thrown one at the time into the boxes.

If the first piece 1 falls in the blue box:

1 ..

The second piece can fall in one of the two boxes and there are two possibilities

12 .. 1 2

The third piece can also follow in one of the two boxes

123 .. 12 3 13 2 1 23

Similarly if the first piece had fallen into the red box we would have obtained the following sequences:

23 1 2 13 3 12 .. 123

We have M^N combinations because each one of the N pieces has M boxes in which it can go. In the previous example we found 8 different combinations, that is 2^3.

If the pieces are not distinguishable so not all those sequences are distinct. In our case the pieces of density are not distinguishable every piece of density is equivalent to any another. In the 8 previous sequences for example: 12 3 ; 13 2 ; 23 1 represents the same state where in the blue box there are two pieces and one in the red. We have taken 2 objects out of 3 and put them into the first box. The number of sequences that represent the same state is

computed as previously shown as $^3C_2=3$ eq.1-86 that gives the number of ways we can pick *m* object from *n* different ones without looking at the order in which they are picked.

We only have 4 different sequences

xxx .. xx x x xx .. xxx

There is one way to obtain the sequence xxx .. or .. xxx, and 3 different ways to obtain xx x or x xx. It will then be more probable to find one of these last two sequences than one where all the pieces are in the same box. The expectation frequency will be 1/8 for the sequence xxx .. and 3/8 for the sequence xx x.

As seen before we have divided our electron density into N pieces and we have also divided the cell into M boxes and the number of pieces N is very large; N>>M. Suppose that after having scattered all these pieces we find n_1 pieces in the first box, n_2 in the second box and so on, then the total number of pieces is defined by:

$$N = \sum_{i=1}^{M} n_i \qquad (1.40)$$

This distribution of $\{n_i\}$is a possible candidate for the density. Since every piece of density can land in any of the boxes, because they are randomly scattered, there are M^N different way to scatter the pieces amongst them and each is as likely to occur. Not all the sequences are distinct as many yield the same distribution as we have just seen in the example of the two boxes and the three pieces. If we find in the first box n_1 pieces, as seen before there are $^NC_{n1}$ different ways to put n_1 pieces out of the total number N in the first box. We are left now with N-n_1 pieces. We find then n_2 pieces in the second box. There are $^{N-n1}C_{n2}$ different ways to put n_2 pieces out of the number N-n_1 of pieces left and so on. When we find a distribution $\{n_i\}$, that means that we find n_1 pieces in the first box, n_2 in the second box and so on, we have then that the number of ways of obtaining this distribution is:

$$^NC_{n_1} \times {}^{N-n_1}C_{n_2} \times {}^{N-n_1-n_2}C_{n_3} \times {}^{N-n_1-n_2-n_3}C_{n_3} \times \ldots\ldots \times {}^{n_M}C_{n_M} = \frac{N!}{n_1!\,n_2!\,n_3!\ldots\ldots n_M!} \qquad (1.41)$$

The expected frequency distribution $F(\{mi\})$ will then be

$$F(\{n_i\}) = \frac{number_of_ways_of_obtaining_a_particular_distribution_n_i}{number_of_possible_way_in_which_the_pieces_can_be_scattered} \quad (1.42)$$

$$number_of_ways_of_obtaining_a_particular_distribution_n_i = \frac{N!}{n_1!\,n_2!\,n_3!.....n_M!}$$

$$number_of_possible_ways_in_which_the_pieces_can_be_scattered = M^N$$

Taking the logaritm of eq. (1.42) we obtain

$$\log(F) = -N\log(M) + \log(N!) - \sum_{i=1}^{M}\log(n_i!) \quad (1.43)$$

We can suppose to divide our density into a huge number of small pieces and so we have that $n_i \to \infty$. In this case we can use the Stirling approximation :

$$\log(n!) \cong n\log(n) - n \quad (1.44)$$

and eq.(1.43) will became:

$$\log(F) = -N\log(M) + N\log(N) - \sum_{i=1}^{M} n_i\log(n_i) \quad (1.45)$$

Since the log is a monotonic function, the maximum of eq.(1.45) will give us the maximum of the expected frequency distribution that is the assignment that best represents our state of knowledge. As we have already said: after many trials, some distribution $F(\{mi\})$ will be found to come up more often than others; the one that most frequently occurs and is consistent with the observed data will be a sensible choice for the density.

As N and M are constant the maximum of log(F) will be equivalent to computing the maximum of:

$$S = -\sum_{i=1}^{M} n_i\log n_i \quad (1.46)$$

This function S is called entropy.

Up to now we have considered all the boxes as having the same dimension. If the boxes have different dimensions there is a different probability to be occupied, bigger as bigger is the dimension of the box. For each box then we define a probability m_i to be occupied and the sum over all the boxes of the m_i probabilities is equal to one:

$$\sum_{i=1}^{M} m_i = 1 \tag{1.47}$$

In this case then eq.(1.42) will became:

$$F(\{n_i\}) = \frac{N!}{n_1!\,n_2!\ldots\ldots n_M} \times m_1^{n_1} m_2^{n_2} \ldots\ldots m_M^{n_M} \tag{1.48}$$

if all the probabilities are the same we obtain again the term $1/M^N$.

Taking the logarithm and using Stirling's approximation we have that:

$$\log(F) = N\sum_{i=1}^{M} n_i \log(m_i) - N\sum_{i=1}^{M} n_i \log(n_i) \tag{1.49}$$

We have now that the entropy in a more general way:

$$S = -\sum_{i=1}^{M} n_i \log(\frac{n_i}{m_i}) \tag{1.50}$$

Let's now see the application of all this.

We have seen that the distribution $F(\{ni\})$ that most frequently occurs is the one that maximises the entropy function and is consistent with the observed data.

Let us suppose that we only know the normalization condition, then the number of electrons computed by the MEM must be equal to the actual number of electrons presents in the unit cell. The normalization condition is expressed in mathematical term as:

$$C_N = N_{el} - \frac{V_{uc}}{N_{pix}} \sum_{i=1}^{N_{pix}} \rho_i = 0 \tag{1.51}$$

where N_{pix} is the number of pixels into which the unit cell has been divided, V_{uc} the volume of the unit cell and ρ_i is the electron density of the ith-pixel. We have then to maximize the entropy subject to this constraint. Maximizing a function subject to a constraint is done using the method of Lagrange undetermined multipliers; this is equivalent to unconstrained maximization with respect to ρ and λ_N of the Lagrangian:

$$
\begin{aligned}
&Q(\rho) = S(\rho) - \lambda_N C_N \\
&\frac{\partial Q}{\partial \rho_i} = \frac{\partial S}{\partial \rho_i} - \lambda_N \\
&\frac{\partial Q}{\partial \rho_i} = -1 - \log \rho_i - \lambda_N = 0 \\
&\rho_i = \exp(-1 - \lambda_N) \\
&\sum_{i=1}^{N_{pix}} \rho_i = \frac{N_{el} N_{pix}}{V_{uc}} \\
&N_{pix} \exp(-1 - \lambda_N) = \frac{N_{el} N_{pix}}{V_{uc}} \\
&\rho_i = \frac{N_{el}}{V_{uc}}
\end{aligned}
\tag{1.52}
$$

We have predicted a uniform density ρ. This is appropriate because we have a complete ignorance about our density and in fact we have started the computation with only the knowledge of the normalization condition.

If for example we have 314 electrons in a cell of 1000 Å^3 and we divide the cell into 20000 pixels, we obtain the value of ρ_i = 0.314 e/Å^3 for every pixel. As the pixel has a volume of V_{UC}/N_{pix}= 0.05 Å^3, in every pixel 0.0157 electrons are located.

If we have measured some experimental data we have a new condition to be satisfied:

$$
C_F = \frac{1}{N_F} \sum_{i=1}^{N_F} \left(\frac{| F_{obs}(\vec{H}_i) - F_{MEM}(\vec{H}_i)}{\sigma(\vec{H}_i)} \right)^2 - 1 = 0 \tag{1.53}
$$

where N_F is the number of observed structure factors.

The maximum entropy criterion selects only from the solutions that fit the data that are supposed to be noisy with a Gaussian distribution of the errors.

Now the Lagarangian is:

$$
Q(\rho) = S(\rho) - \lambda_N C_N - \lambda_F C_F \tag{1.54}
$$

using the more general definition of entropy as in eq.(1.50) $S = -\sum_{i=1}^{M} \rho_i \log(\frac{\rho_i}{\tau_i})$ we obtain:

$$
\begin{aligned}
&\frac{\partial Q}{\partial \rho_i} = \frac{\partial S}{\partial \rho_i} - \lambda_F \frac{\partial C_F}{\partial \rho_i} - \lambda_N = 0 \\
&-1 - \ln \frac{\rho_i}{\tau_i} - \lambda_F \frac{\partial C_F}{\partial \rho_i} - \lambda_N = 0
\end{aligned}
\tag{1.55}
$$

This gives an implicit solution for the MaxEnt electron density ρ_{MEM} in the form:

$$\rho_i = \tau_i \exp(-1-\lambda_N)\exp(-\lambda_F \frac{\partial C_F}{\partial \rho_i}) \tag{1.56}$$

Using the normalization condition of eq.(1.51) we obtain:

$$\sum_{i=1}^{N_{pix}} \tau_i \exp(-1-\lambda_N)\exp(-\lambda_F \frac{\partial C_F}{\partial \rho_i}) = \frac{N_{el}N_{pix}}{V_{uc}} \tag{1.57}$$

combining these two last equations eq(1.56) and eq(1.57) ,we can eliminate the Lagarange multiplier λ_N:

$$\rho_i = \frac{N_{el}N_{pix}}{V_{uc}} \tau_i \exp(-\lambda_F \frac{\partial C_F}{\partial \rho_i}) / \sum_{j=1}^{Npix} \tau_j \exp(-\lambda_F \frac{\partial CF}{\partial \rho_j}) \tag{1.58}$$

This equation can't be solved analytically as C_F depends on ρ_i:

$$F_{MEM}(\vec{H}) = V\sum_l \rho(\vec{l})\exp(-2\pi i\vec{H}\cdot\vec{l}) \tag{1.59}$$

Several algorithms have been developed, that solve the problem iteratively. We will concentrate here only on the one that is used in the program PRIMA[39] that we used in this thesis to compute the MEM maps.

1.2.6.1 PRIMA algorithm

To solve eq.(1.58) we need to introduce approximations. Kumazawa introduced the zero-order single-pixel approximation (ZSPA)[40], where we expand the derivative of C_F with respect to ρ_i in a Taylor series:

$$\frac{\partial C_F}{\partial \rho_i} = \frac{\partial C_F}{\partial \tau_i} + \sum_{l=1}^{Npix} (\rho_l - \tau_i)\frac{\partial^2 C_F}{\partial \tau_i \partial \tau_l} + ... \tag{1.60}$$

and we neglect all the terms except the first one. This linear approximation is equivalent to the simple replacement of $\frac{\partial C_F}{\partial \rho_i}$ with $\frac{\partial C_F}{\partial \tau_i}$. The second approximation is:

$$\tau_i^{(n+1)} \approx \rho_i^{(n)} \tag{1.61}$$

This means that the MEM electron density of cycle n is used as prior electron density for cycle n+1.

With these approximations the formula for the iteration is:

$$\rho_i^{(n+1)} = \frac{N_{el}N_{pix}}{V_{uc}} \rho_i^{(n)} \exp(-\lambda_F \frac{\partial C_F}{\partial \rho_i}\Big|_{\rho_i^{(n)}}) / \sum_{j=1}^{Npix} \rho_j^{(n)} \exp(-\lambda_F \frac{\partial C_F}{\partial \rho_j}\Big|_{\rho_j^{(n)}}) \qquad (1.62)$$

The iteration is started with $\rho_i^{(1)} = \tau_i$

We can make a choice between two kinds of prior information about the electron density distribution. One is a uniform density in the unit cell, which correspond to the maximum-entropy state where there is no structural constraint. The other choice is to use density computed by conventional Fourier series.

The last problem is to determine the value of λ_F . It has been shown that the result of the (ZSPA) is not sensitive to the exact value of λ_F, if it is small enough to assure the convergence for the iteration.

Information about overlapped reflection in powder diffraction data can be introduced into the MEM analysis by using the G-constraint:

$$C_G = \sum_{j=1}^{M_G} \frac{| G_{cj} - G_{oj} |^2}{\sigma_j^2} \qquad (1.63)$$

where

$$G_{cj} = \frac{\sum_{n=1}^{L_j} m(\vec{h}_n) \mid Fc(\vec{h}_n) \mid^2}{\sum_{n=1}^{L_j} m(\vec{h}_n)} \qquad (1.64)$$

In eq.(1.63) M_G is the total number of groups comprising overlapped reflections, G_{oj} is the sum of integrated intensities of overlapped reflections in group j, L_j is the number of overlapped reflections in group j and $m(h_n)$is the multiplicity of the *n*th reflection.

MEM gives a map of higher resolution as compared to a Fourier map. But one has also to take care in interpreting the MEM when looking at the finest details, as we have seen that in the MEM method, some approximations have been made and there is a series termination effect on the Taylor expansion of the constraint derivative C_F that can cause artefacts in the map. A sufficient informative, non-uniform prior electron density can suppress these artefacts almost completely. We will not enter here into the details[41, 42]: the best prior that can be obtained from standard structure refinement is the independent atom model where

analytical Fourier transforms are used and termination errors are avoided, but since it relies on the knowledge of the temperature factors, which are often inaccurate, it can be used only for simple structures. Other improvements can be obtained using generalized F-constraint and prior-derived F-constraint, where in the latter in addition the structure factors that are not known from the experiment will be constrained to values close to F_{prior}, computed from the model, which is the best estimate that we have.

Chapter 2 Instrumentation

Dans ce chapitre, nous présentons succintement les principes de fonctionnement des sources de rayonnement synchrotron ainsi que les diffractomètres utilisés au cours de cette étude. Nous décrivons notamment les quatre lignes (BM16, ID31, SNBL (BM1A) et ID11) de lumière synchrotron de l'ESRF (Grenoble, France) sur lesquelles ont été effectuées les mesures. Une importance particulière est portée au diffractomètre sur poudres haute résolution ID31 avec lequel la plupart des expériences ont été entreprises.

2.1 Synchrotron sources

Synchrotron radiation is the electromagnetic radiation emitted by charged particles moving in circular orbits with highly relativistic velocities. The emission – also called "magnetic Bremsstrahlung" – covers a wide range of wavelengths (10^3 to 10^{-1} Å) and was first detected as the undesirable by-product from high energy particle accelerators[43].

In a modern synchrotron radiation source electrons (or positrons) are accelerated at relativistic speeds and, with the help of dipole bending magnets, they are forced to circulate in a closed orbit inside the vacuum chamber of a storage ring. The energy lost by the particles due to emission of radiation is restored periodically by a radio-frequency cavity system. The synchrotron radiation is emitted from the bending magnets in a broad fan tangential to the particles' trajectory.

An outline of the European Synchrotron Radiation Facility (ESRF), a dedicated third generation source, is shown in Figure 2-1.[44]

In this machine, high-brilliance X-rays are produced by accelerating electrons in a linear accelerator pre-injector (linac) and then in a booster synchrotron. When the particles have reached a high enough energy, (6 GeV), they are transferred to a storage ring equipped with insertion devices and bending magnets. The insertion devices consist of periodic arrays of magnets which modify the trajectory of the electrons in the straight sections of the ring. The

modification of the trajectory from a straight line results in the emission of synchrotron radiation with particular characteristics. In the most common type of insertion devices, wigglers and undulators, the array of magnets is usually designed to produce an approximately sinusoidal magnetic field and, therefore, the resulting particle trajectory is also sinusoidal.

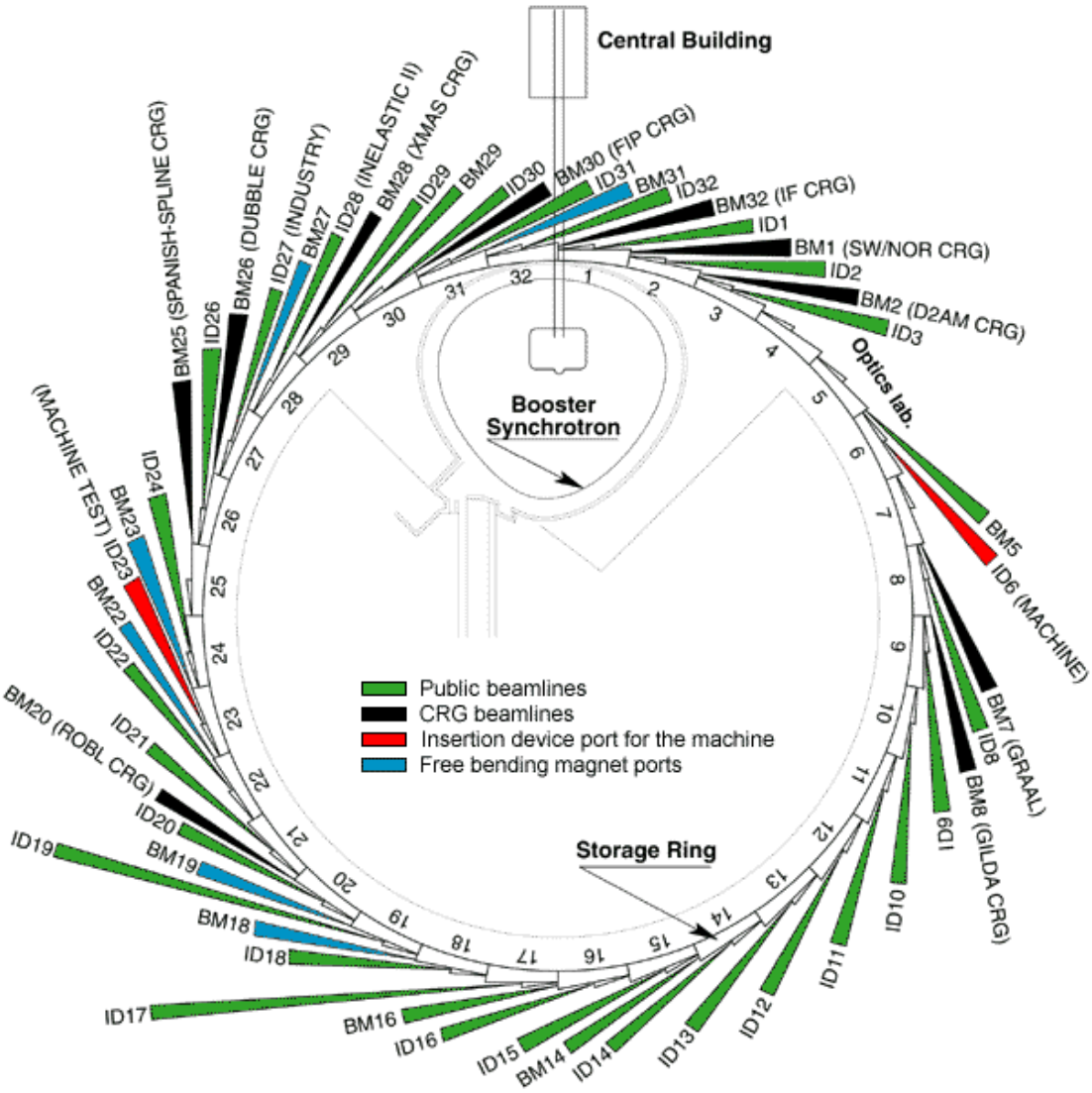


Figure 2-1 Outline of the ESRF at Grenoble

Operating with a strong magnetic field, the synchrotron radiation emitted will be the contribution from each period of the trajectory with negligible interference effects, and therefore, the radiation is equivalent to that of a bending magnet, but multiplied by twice the number of periods in the magnetic array. Operating with weak magnetic field, interference effects are important, producing radiation which is concentrated around one or a few wavelengths and highly collimated horizontally.

In Figure 2-2 is shown the characteristic radiation emitted from a bending magnet, a wiggler and an undulator at ESRF.

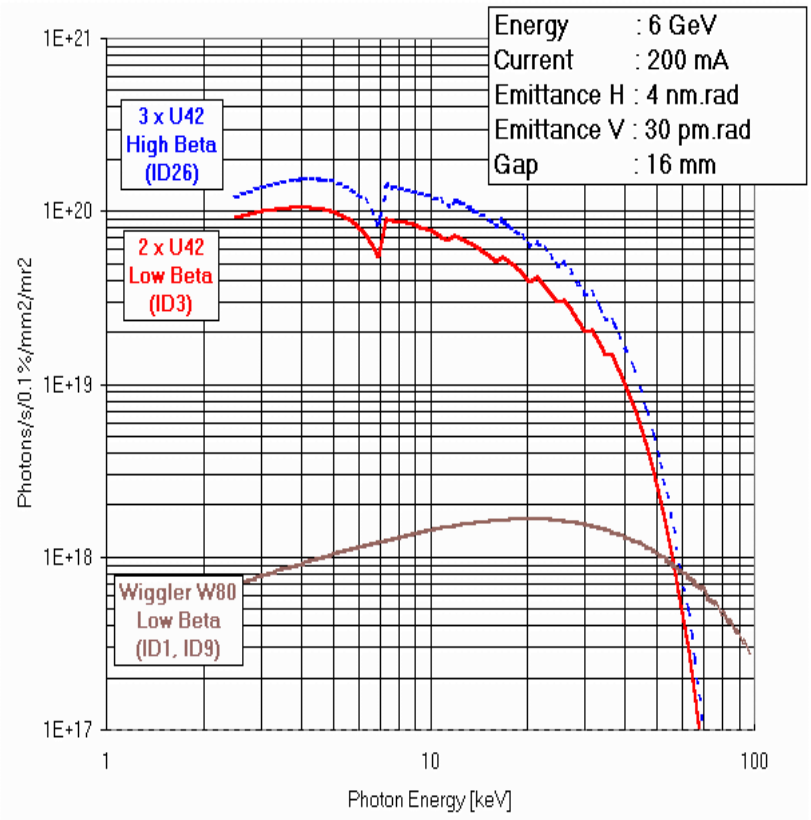


Figure 2-2 Example of brilliance achieved at the ESRFfor different types of insertion device.

The spectral distribution of a wiggler is shifted towards shorter wavelengths and the intensity greatly increased with respect to that of a bending magnet. The radiation from an undulator is very bright and concentrated in a narrow band around the fundamental wavelength. Hence, to a large extent, the radiation can be tailored to the needs of the experiment. Some of the outstanding proprieties of synchrotron radiation are summarised below:

High intensity (brilliance)

Continuous spectrum from the UV to the X-ray region.

Very low beam divergence, especially, in the vertical plane.

Linear polarization of the radiation, with the electric vector parallel to the plane of the orbit.

Time structure of the beam, with the length of the pulses varying between 2 *ns* to 1000 *ns*.

These proprieties, especially allow the design of versatile and highly accurate instrumentation, and two types of powder diffraction experiments are available, those which use monochromatic radiation and those which use the continuous spectrum.

In the monochromatic experiment, a specific wavelength can be selected from the white beam using a double crystal monochromator. Because of the high intensity and the low divergence, the angular resolution is far superior to what is attainable using a conventional X-ray tube. This results in a significant reduction in peak overlap and therefore, this type of data is useful for direct solution and refinement of structures. Furthermore, indexing the diffraction pattern is also facilitated, because of the better separation between reflections (especially at low angle) and the accuracy in the determination of the position of the peaks.

2.2 High resolution powder diffraction beamline, ID31/BM16

The beamline BM16 was built on a bending magnet and designed to operate in the energy range 5 - 40 keV. The diffractometer was moved to the new beamline ID31 in March 2002. BM16 could accept 4 mrad of white X-ray radiation in the horizontal plane from the bending magnet. This was incident on a curved mirror, set at grazing incidence, which collimated the beam in the vertical sense. After the mirror, the residual vertical divergence was around 12 μrad FWHM. It was this high degree of vertical collimation that was ultimately responsible for the high angular and energy resolution of the beam line. After the mirror the beam was monochromated by a water-cooled double-crystal monochromator using Si 111 reflections. Bending of the second crystal sagittally gave the option to focus the beam horizontally onto the sample. A second mirror also gave the option to focus the beam vertically onto the sample. Both horizontal and vertical focussing led to a loss of angular resolution, for they increased the vertical divergence of the beam incident on the sample, so these options were rarely exploited.

The new beamline ID31 has three undulators with a minimum gap of 11 mm, allowing the energy range from 5 keV to 60 keV (0.21 Å) to be covered without interruption. No optical elements are used other than a liquid-nitrogen-cooled double-crystal mono-chromator with a choice of Si 111 or Si 311 crystals. The undulator beam is highly collimated in the horizontal and the vertical directions. The axial dimension of the beam on the sample is from 1 mm to 2 mm, and the receiving slits are from 1 mm to 2 mm, thus reducing the marked asymmetry at low angle seen on BM16 along with the associated peak broadening and peak shifts. The diffractometer, the same one as on BM16, is mechanically very stable, accurate (±1 arcsec) and precise (sensitive to displacements of 5×10^{-5}°). It can accept spinning capillary or flat plate specimens. In routine operation, a bank of nine detectors is scanned to measure the diffracted intensity as a function of 2θ (Figure 2-3). The detectors are ≈2° apart, and each is preceded by a Si 111 analyser crystal. The nine crystals are scanned together, as a unit, effectively measuring nine high resolution diffraction patterns in parallel, with 2° between them.[45]

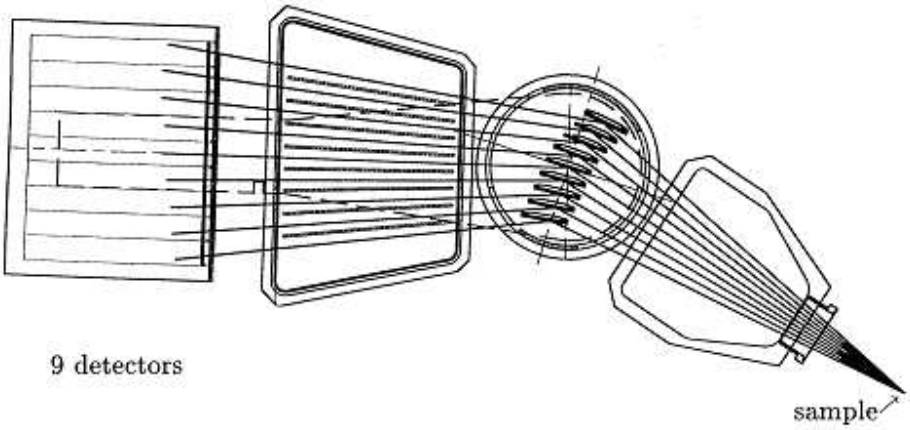


Figure 2-3 Diagram of the nine-channel multianalyser detector

2.3 The Swiss-Norwegian beamlines, SNBL (BM1A)

The beam line is built on a bending magnet and is designed to operate in the energy range 5 - 40 keV. The incoming horizontal fan of radiation is first divided into two separate beams. In the case of the BM1A station, a 2 mrad horizontal fan of radiation is first vertically collimated by a primary mirror. The first crystal of the monochromator is water cooled, while the second crystal can be sagittally bent for horizontal focusing. A variable gap between the two crystals, combined with a translation of the second crystal along the beam direction, allows for a fixed exit beam height. An image-plate area detector (MAR) is available for macromolecular crystallography and powder diffraction work. This is a 345mm diameter image plate with online readout capabilities. The pixel resolution is either 0.1 mm or 0.15 mm, and four different readout sizes can be chosen (180, 240, 300 and 345 mm). For example, choosing readout sizes 345 mm or 240 mm, with a fixed pixel resolution of 0.15 mm, images of 2300 or of 1600 pixels, respectively, can be obtained. Readout times (scan +

erase) vary between 35 sec and 110 sec. The focal spot achieved with the combination of mirror and crystal focussing has a diameter of about 500 μm.

2.4 The materials science beamline, ID11

ID11 is a multipurpose beamline concentrating on experiments in the area of materials science. The experimental hutch dedicated to diffraction experiments has an extremely high flux, even at high energies, making it suitable for time-resolved experiments, diffraction from microcrystals and for diffraction from crystals in absorbing sample chambers such as furnaces, high-pressure cells or cryostats. The wavelength is tunable in the energy range 7-100 keV with an energy resolution close to $\Delta E/E = 10^{-4}$ up to 40 keV.

The main optical elements are a 1 m long Pt-coated Si premirror, a double crystal monochromator, and a second 1.2 m long Pt-coated mirror. Other components are an attenuator system in front of the first mirror and water-cooled slit systems of two types: one primary slit system for beam definition, and four secondary slit systems for beam conditioning.

We made use of a Bruker Smart 1000 CCD detector system to collect 2 dimensional powder diffraction rings.

Chapter 3 The low temperature phases of globular organic molecules

Ce chapitre est consacré aux phénomènes d'ordre-désordre observés dans certains solides moléculaires organiques de type globulaire. Ces molécules de forme presque sphérique possèdent en effet souvent une phase solide à haute température. Dans ce cas, les cristaux sont facilement déformés mécaniquement, et présentent des structures cristallines de haute symétrie dont l'ordre augmente lorsque la température décroît. Ces structures étaient auparavant inconnues, et nous les avons déterminées en utilisant des simulations de Monte-Carlo. Nous expliciterons en premier lieu les caractéristiques des cristaux moléculaires, puis nous exposerons les résultats experimentaux obtenus sur divers terpenes.

3.1 Molecular crystals

Molecular crystals[1, 46] are formed by discrete molecular entities packed together by weak dispersion or polar forces. Interatomic distances within the crystal can be generally grouped into two distinct distributions, firstly of bond distances within the molecules and secondly of much longer contact distances between atoms of different molecules. The forces determining the packing in molecular crystals are much weaker than those of ionic or covalent chemical bonds. The lattice energy U varies from -0.5 to -4 kcal mol^{-1} for noble gases up to -40 to -50 kcal mol^{-1} for neutral molecules of significant complexity. On the other hand the lattice energy of a covalent compound such as SiO_2, for example, is -3137 kcal mol^{-1}. Such weaker intermolecular forces can be approximately classified as follow:

Repulsive forces

Most stable molecules have closed electron shells and cannot accept other electrons without violating the Pauli exclusion principle. They repel other molecules trying to come close. Repulsive forces exert their effect at very short range, approximately decreasing with the

twelfth power of the interatomic distance and grow rapidly when two non-bonded atoms try to get closer than their proper contact distance, given by the sum of their van-der-Waals radii.

Dipole-dipole interactions

Molecules having permanents dipole moments experience electrostatic attraction when properly oriented. Another effect to consider is that a polar molecule can induce a dipole in a neighbouring polarizable molecule. It can be shown that the interaction energy depends on the inverse sixth power of the distance between the two dipoles.

Dispersion or London forces

Non-polar molecules attract one another even though neither has a permanent dipole moment. The evidence for the existence of interactions between them is the formation of condensed phases of nonpolar substances, such as the condensation of hydrogen or argon to a liquid at low temperature. The interaction between nonpolar molecules arises from the transient dipole that all molecules posses as a result of fluctuations in the instantaneous positions of electrons. If the electrons in one molecule fluctuate into an arrangement that gives the molecule an instantaneous dipole moment μ_1, this dipole will generate an electric field that polarizes another molecule and will induce in that molecule an instantaneous dipole μ_2. The two dipoles attract each other and the potential energy of the pair is lowered.

Hydrogen bonding

The hydrogen bond is the most important of all directional intermolecular interactions[47]. It is operative in determining molecular conformation, molecular aggregation, and the function of a vast number of chemical systems ranging from inorganic to biological. There are dozens of different types of $X-H\cdots A$ hydrogen bonds that commonly occur in the condensed phases, and in addition there are numerous less common ones. Dissociation energies span more than two orders of magnitude (about 0.2-40 kcal mol^{-1}). Within this range the nature of the interaction is not constant, but electrostatic, covalent and dispersion contributions vary in

their relative weights. The hydrogen bond has a broad transition region that merges continuously with covalent bonds, van-der-Waals interactions, ionic interactions and also cation -π interactions.

In a hydrogen bond $X-H\cdots A$, the group $X-H$ is called the donor and A is called the acceptor (short for "proton donor" and "proton acceptor"). Since the hydrogen bond has a long range, a donor can interact with two or three acceptors simultaneously.

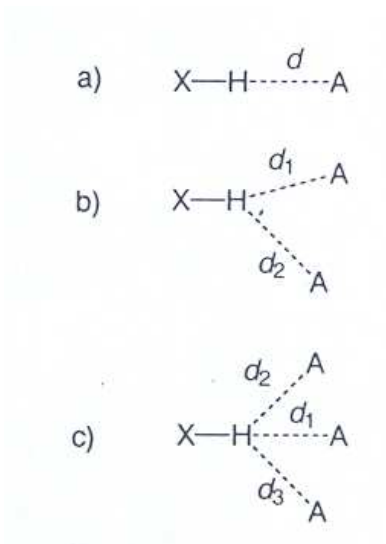


Figure 3-1 Different types of hydrogen bridges

The energy of hydrogen bonds in solids cannot be directly measured, and this circumstance leaves open questions in many structural studies. In theoretical studies energy values are computed for dimers in optimal geometry unperturbed by their surroundings. In the solid state, hydrogen bonds are practically never in optimal geometry, and are always influenced by their environment. There are numerous effects from the near and also from the remote surroundings that may considerably increase or lower hydrogen bond energies ("crystal field effects"). Hydrogen bonds do not normally occur as isolated entities but form networks. Within these networks, hydrogen bond energies are not additive. The total energy of a hydrogen bond can be split into contributions from electrostatic, polarization, charge

transfer, dispersion and exchange repulsion. Chemical variation of donor and/or acceptor, and possibly also of the environment, can gradually change a hydrogen bond to another interaction type.

The transition to a pure van der Waals interaction is very common. The polarity of $X-H$ or A in the array $X^{\delta-}-H^{\delta+}\cdots A^{\delta-}$ can be reduced by suitable variation of X or A. This reduces the electrostatic part of the interaction, whereas the van-der-Waals component is much less affected. In consequence, the van-der-Waals component gains relative weight, and the angular characteristic gradually change from directional to isotropic. Since the polarities of $X^{\delta-}-H^{\delta+}\cdots A^{\delta-}$ can be reduced to zero continuously, the resulting transition of the interaction from hydrogen bond to van der Waals type is continuous too. Such behaviour, for example, is demonstrated for the directionality of C-H···O=C interactions which gradually disappears when the donor is varied from C≡C-H to C=CH_2 to C-CH_3

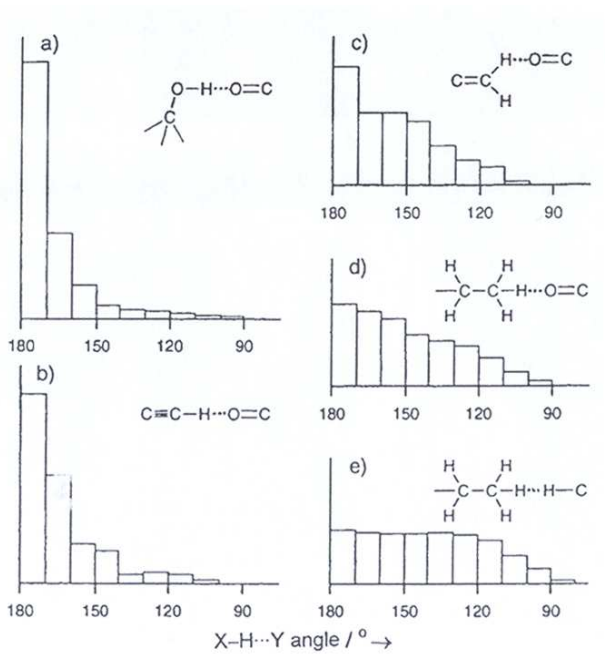


Figure 3-2 Directionality of X-H···O=C interactions with X-H groups of different polarities

There is also a continuous transition to covalent bonding. In the so called symmetric hydrogen bond X-H-X, where an H atom is equally shared between two chemically identical atoms, no distinction can be made between a donor and an acceptor, or a "covalent" X-H and a "noncovalent" H···X bond. In crystals all the intermediate cases exist between the extremes X-H-X and X-H·····X.

There is a gradual transition from hydrogen bonding to purely ionic interactions. If in an interaction $X^{\delta-} - H^{\delta+} \cdots Y^{\delta-} - H^{\delta+}$ the net charges are zero, the electrostatics are of the dipole-dipole type. In general however the net charges are not zero. Alcoholic O-H groups have a partial negative charge in addition to their dipole moment; the ammonium group has a positive net charge and so on. This situation leads to ionic interaction between the charge centres with the energy having an r^{-1} distance dependence. If the charges are large the ionic behaviour may became dominant.

Finally there is a transition region between the hydrogen bond and the cation-π interaction. In the pure cation-π interaction a spherical cation such as K^+ contacts the negative charge concentration of a π-bonded moiety such as a phenyl ring. This can be considered an electrostatic monopole-quadrupole interaction. The bond energy is -19.2 kcal mol^{-1} for the example of K^+····benzene. A pure π-type hydrogen bond $X^{\delta-} - H^{\delta+}$ ····Ph is formally a monopole-quadrupole interaction with much lower energies of only a few kcal/mol^{-1}. If charged hydrogen bond donors such as NH_4^+ interact with π-electrons clouds, local dipoles are oriented at the π face, but the energetics are dominated by the charge-quadrupole interaction.

3.1.1 Crystal packing

In a crystal relative molecular positions are mainly determined by short range forces. Calling r the interatomic distance, the repulsion energy drops nearly with r^{-12}, the attractive dispersion energy tends to zero with r^{-6}, and the dipole-dipole and monopole-monopole attractive

electrostatic energies decrease with r^{-3} and r^{-1}, respectively. Since van der Waals forces are both adirectional and additive, the energy of any single atom is lower the higher the number of atoms of other molecules surrounding it at contact distances. The most stable crystal is the one in which the molecules pack themselves with the highest coordination number. An exception to this rule is represented by the hydrogen bond, that requires directional interactions. A geometrical analysis has been carried out by Kitaigorodsky who has shown that rows of molecules staggered by a glide lattice operation can produce a very efficient close packing by repeating a molecule of arbitrary form in the space group P1, $P2_1$, Pc, $Pca2_1$, $Pna2_1$ and $P2_12_12_1$ or a centrosymmetric molecule in the space group P-1, $P2_1/c$, C2/c or Pbca. These space groups are the most frequently observed for molecular crystals.

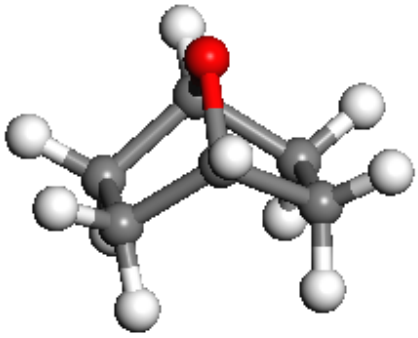
3.2 Plastic crystals

Crystals with unusual mobility were first recognized by Timmermans[48]. These crystals, called plastic crystals, are composed of almost spherical molecules ("globular") and they have unusually low entropy of fusion, the value being of the order of *R*. This fact indicates that only the centre of mass of the molecules acquired mobility on melting, whereas the rotational degrees of freedom become mobile at lower temperatures. In all such cases the melting point of the compound is much higher than expected, indicating that the crystal is stable in spite of considerable mobility of the constituent molecules. Owing to their globular nature and the absence of strong directional intermolecular forces, these molecules have high-symmetry, orientationally-disordered, plastic phases at ambient temperatures and pressures, frequently with cubic or hexagonal symmetry. The formation of the plastic phase is due to the ability of the constituent molecules, over a particular range of temperature or pressure, to gain enough energy to surmount the rotational energy barrier but not to the extent of breaking the lattice structure. Rotation in plastic crystals can be described in terms of orientation disorder with the molecules vibrating around discrete, distinguishable positions in the crystalline lattice and reorienting with very high frequencies. Hence, the situation is not rotation in the usual sense, but rather a restricted rotation from one preferred orientation to another[49].

At the upper limit of this temperature or pressure range, melting occurs with a breakdown of the cubic or hexagonal arrangement but with only a small increase in entropy and little increase in volume. On cooling, the internal energy of the molecule starts to dominate over the entropy factor and a transition takes place to a more- ordered phase, characterised by a lower symmetry crystal structure and by the presence of long-range orientational order. If the transition leads to a fully ordered phase, then the process is usually accompanied by a large change in entropy. Conventionally, the form stable at the highest temperature is designed I, which on cooling gives a form II, which on further cooling gives a form III, and so on. Usually the phase I is the plastic phase and II and any other low temperature forms are the ordered phases, but examples can be cited of substances, such as the 7-oxabicyclo[2.2.1]heptane and the camphor[50, 51], where there exist two plastic phases I and II, before reaching on cooling the ordered phase III.

Terpenes are widespread in nature, synthesized in plants as constituents of essential oils. Many terpenes are hydrocarbons, but alcohols, aldehydes and ketones (terpenoids) are also found. For the hydrocarbons, the basic building block is 2-methyl-1,3-butadiene, $CH_2=C(CH_3)-CH-CH_2$, following the so-called "isoprene" rule identified by Wallach in 1887 . Terpenes are classified according to the number of isoprene units. Monoterpenes have two isoprene units, sesquiterpenes three, diterpenes four, etc. Terpenes and terpenoids can be cyclic, (e.g. the monoterpenes α- and ß-pinene, camphor ($C_{10}H_{16}O$) and menthol), aliphatic, (e.g. monoterpenoids nerol and geraniol), or contain both straight-chain and cyclic components, (e.g. the tetraterpene provitamin A1, ß-carotene, with 8 isoprene units, $C_{40}H_{56}$). The terpenes and terpenoids form a wide class of compounds, with many isomeric forms as isoprene units can be combined with the double-bonds either cis or trans. Many have biological importance, such as ß-carotene, phytol (vitamin A1), or squalene, a precursor in the synthesis of cholesterol. Many smell pleasant and are used in a various preparations, from bath oil to cough lozenges, skin-rubbing lotions, and "alternative medicine", etc. As seems surprisingly often the case with basically simple compounds that have been part of the general environment, rather little is known about the solid-state behavior.

3.2.1 7-oxabicyclo[2.2.1]heptane



The crystal structures of 7-oxabicyclo[2.2.1]heptane $C_6H_{10}O$ (Aldrich,98%) in the low temperature ordered phases were initially unknown. Studies by Ralph M. Paroli et al.[52] using spectroscopic and differential scanning calorimetric techniques showed that four different solid phases may occur, with hysteresis in the transition temperatures. They have seen phase III appear only on warming the sample slowly, otherwise they could only observe the presence of three phases: phase I, II and IV.

At room temperature the sample is liquid as are a lot of globular molecules. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from the liquid phase to the low temperature phase was investigated by recording on the beamline BM16 short diffraction patterns of 2 min each. To see if hysteresis was present in the transition temperatures we cycled a few times up and down in temperature and also with different warming rates. In the next tables we summarize the results obtained.

Transition	Cooling(DSC*) 20K min^{-1}	Warming(DSC*) 20 K min^{-1}	Cooling(us) 4 K min^{-1}	Warming(us) 0.2 K min^{-1}
Liquid-phase I	244 K	243 K	237 K	251 K
Phase I - II	206 K	231 K	180 K	234 K
Phase II-IV	170 K	191 K	155 K	II-III 191 K III-IV 182K

*Table 3-1 Temperature at which the phase transition occurs, * is the work of Ralph M. Paroli et al.*

Warming (0.2 K min^{-1})	T(K)
Ordered phase IV	< 182
Change	182—185
Ordered phase III	185—191
Change	191—192
Disordered phase II	192—234
Disordered phase I	234—251
Begins to melt	252
Liquid	> 253

Table 3-2 Detailed temperatures at which we observed the phase transitions

The phase III appears only if we warm up slowly and it lasts for an interval of only 5 K. This is in agreement with the DSC data of Paroli et al.

The measurements where first performed on the beamline BM16 and then, as we will show, phase III exhibited a strong preferred orientation effect. We have also measured the sample on the Swiss Norwegian beamline (SNBL) where a MAR image plate detector was used in order to try to overcome this problem.

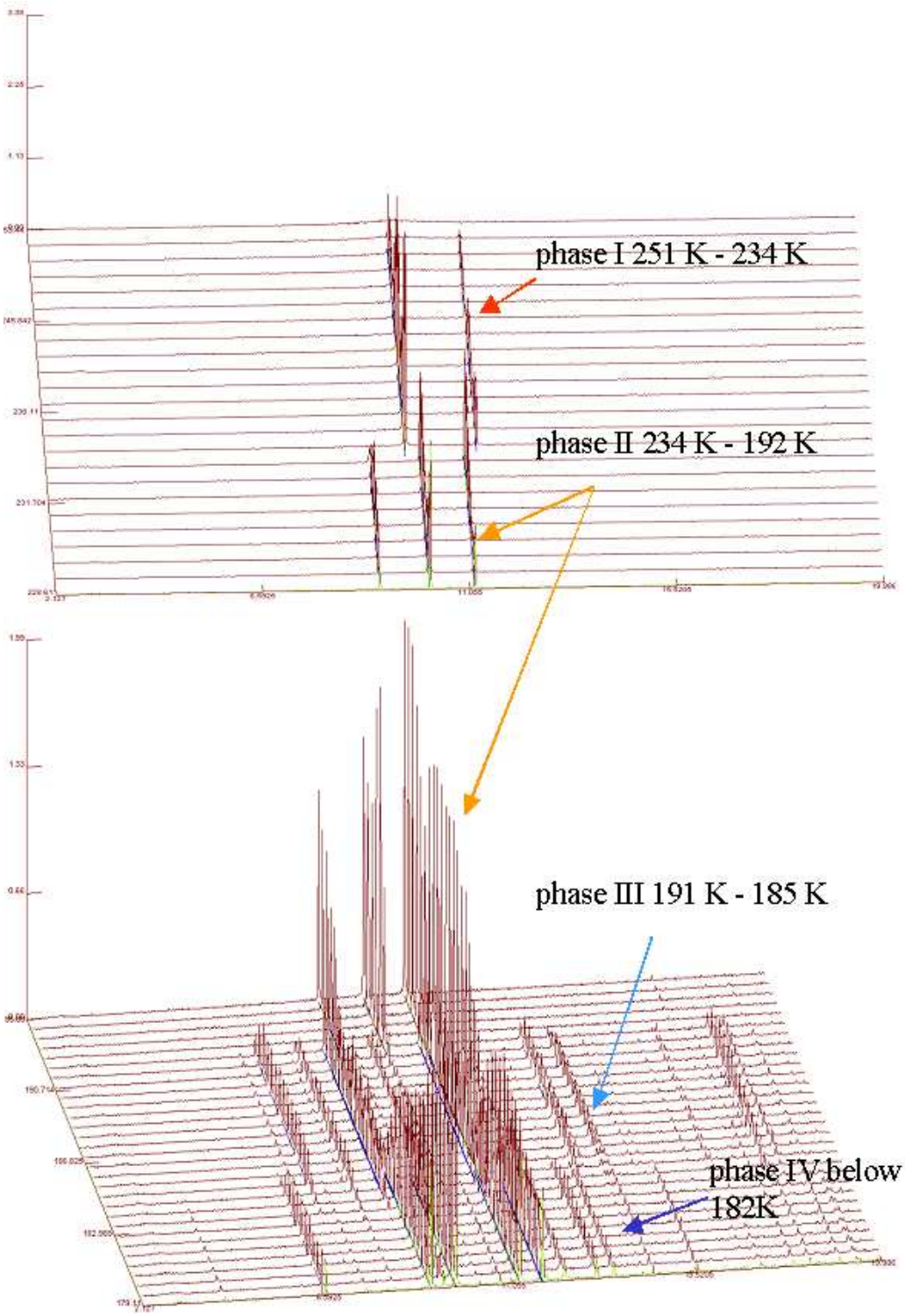


Figure 3-3 Powder diffractograms showing the structural phase transitions

3.2.1.1 Phase I

The lattice parameters of phase I where obtained after the measurements performed on the SNBL. The resolution was worse than on BM16 but the statistics are much higher and we could see some peaks appearing in the low angle region that we could hardly see from the BM16 data. The cell was found with the SVD auto-indexing program present in TOPAS. At 240K the cell is cubic with a = 8.54690(8) Å, space group Fm3m.

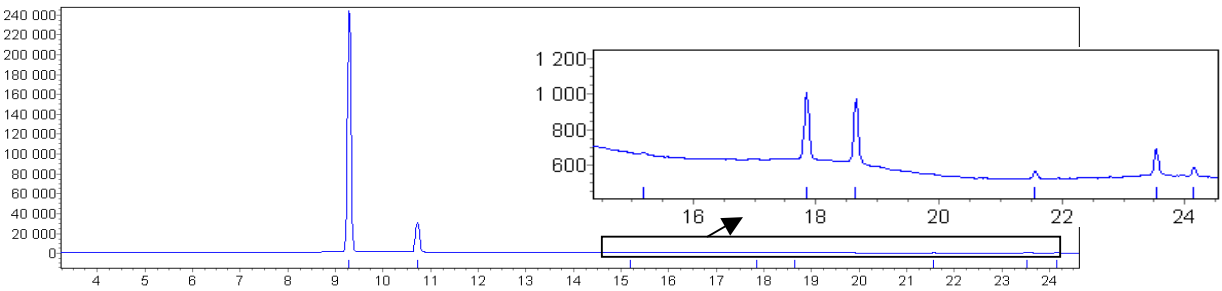


Figure 3-4 Powder diffraction pattern of the disordered face centered cubic phase I (λ=*0.79945* Å)

This behaviour is typical of a plastic crystal. We note also that the sample presents preferred orientation effect. The following image is taken with a capillary rotation around its axis of 80°.

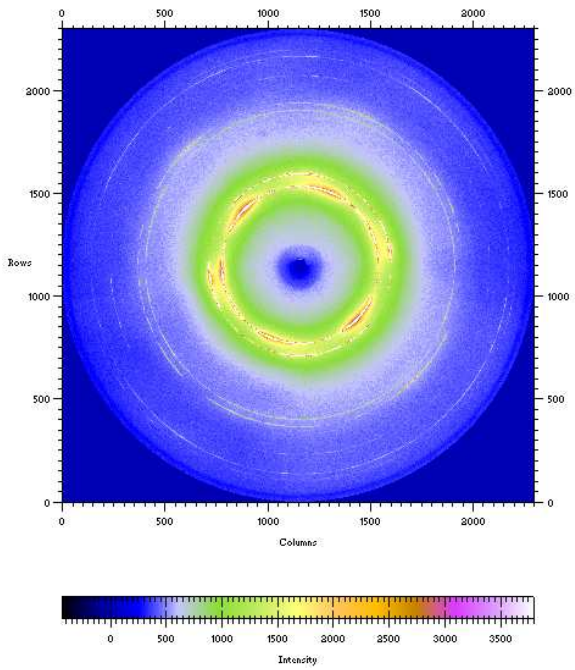


Figure 3-5 The 2-dimensional powder diffraction pattern collected on the MAR image plate on the SNBL.

3.2.1.2 Phase II

This phase was already indexed manually as cubic with a=10.60558(4) Å at T=194 K looking at the BM16 data, but it was difficult to determine the space group as we could only see five peaks. Looking at the data collected on the SNBL, where the number of observed peaks is much higher, it was possible to assign the space group as Pm3m. The cell was checked also with the SVD procedures and the same cell already determined manually was found. The transition is then from phase I, face centred cubic lattice, to phase II, primitive cubic lattice. A few peaks can't be indexed, and are identified as coming from the ice formed on the capillary during the experiment.

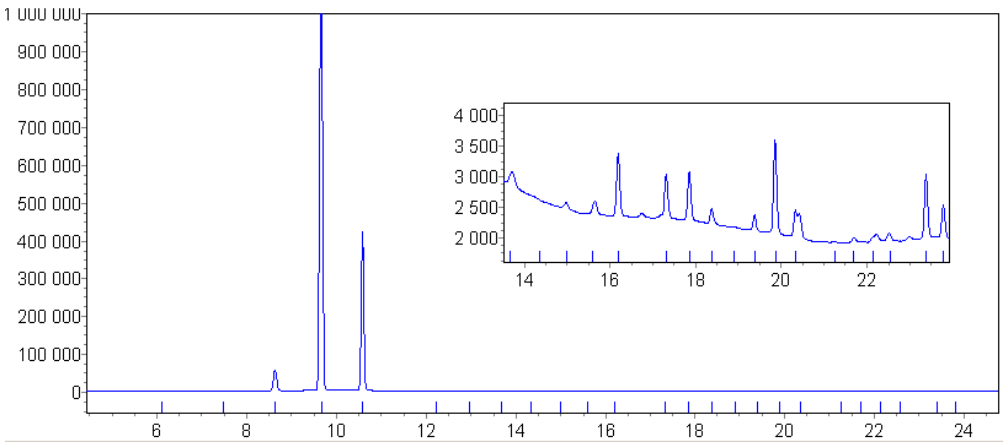


Figure 3-6 Powder diffraction pattern of the disordered primitive cubic phase II (λ=0.79945 Å)

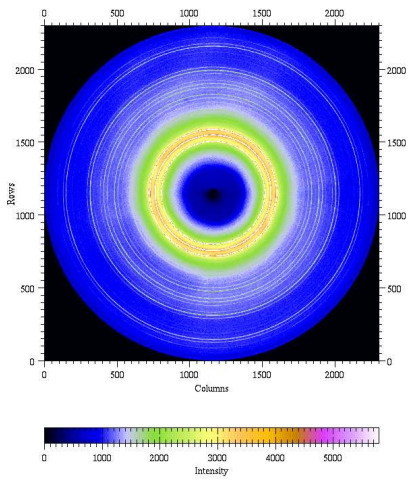


Figure 3-7 The 2-dimensional powder diffraction pattern collected on the MAR image plate on the SNBL.

3.2.1.3 Phase III

The diffraction pattern recorded on the beamline BM16 at 185 K was indexed from the positions of 20 low-angle peaks using the program ITO, with a figure of merit M_{20} of 81, in a monoclinic cell:

$$a = 21.852(1)\ \text{Å}\quad b = 10.9699(4)\ \text{Å}\quad c = 9.3331(4)\ \text{Å}\quad \beta = 90.453(5)^\circ$$

It was not possible initially to solve the structure because of strong granularity but owing to the high resolution it was possible to obtain the unit cell.

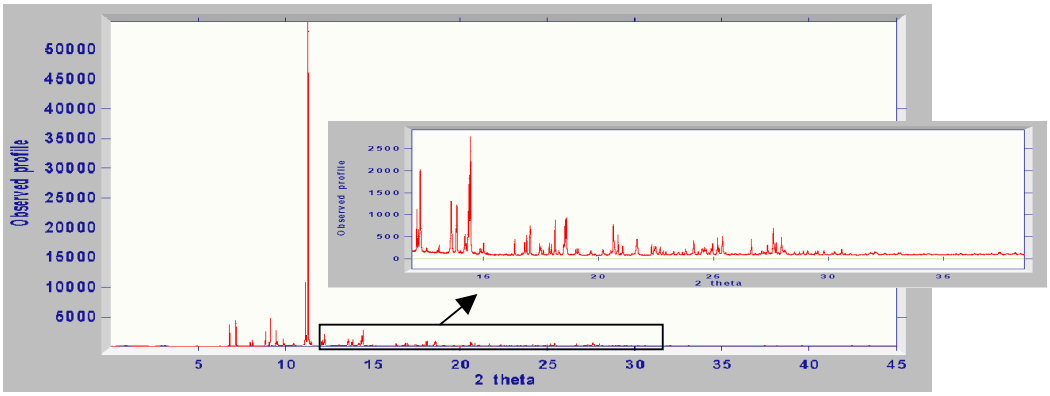


Figure 3-8 The high resolution diffractogram of the phase III collected on BM16.

We performed then another measurement at the same temperature at the SNBL using a MAR image plate that shows the strong granularity present in the sample.

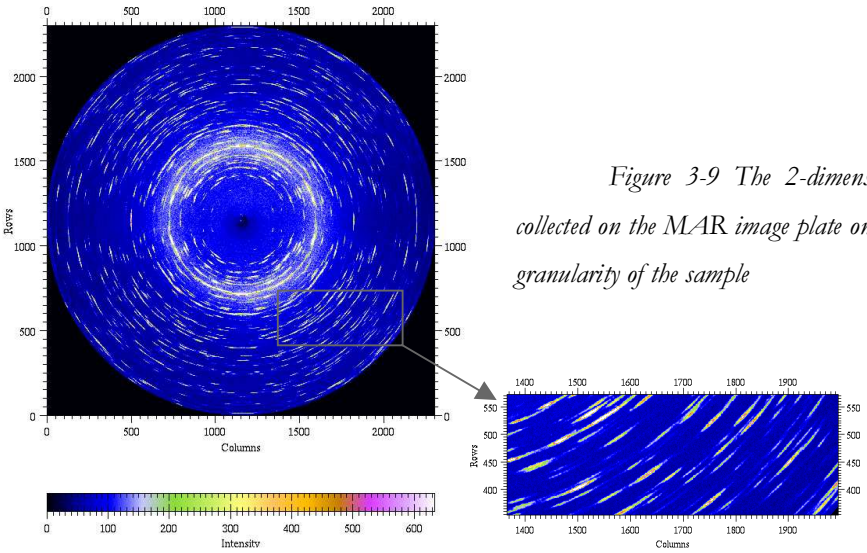


Figure 3-9 The 2-dimensional powder diffraction pattern collected on the MAR image plate on the SNBL that clearly show the granularity of the sample

Averaging the intensities over all the diffraction circles was enough to obtain a dataset that was used to solve the structure with the program FOX (parallel tempering). The cell volume of 2237.3(1) $Å^3$ indicates 16 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_1/a$, and hence there are 4 independent molecules in the asymmetric unit.

In the following picture it is possible to see the different resolution that can be obtained on BM16 and now on ID31 compared to that of an image plate.

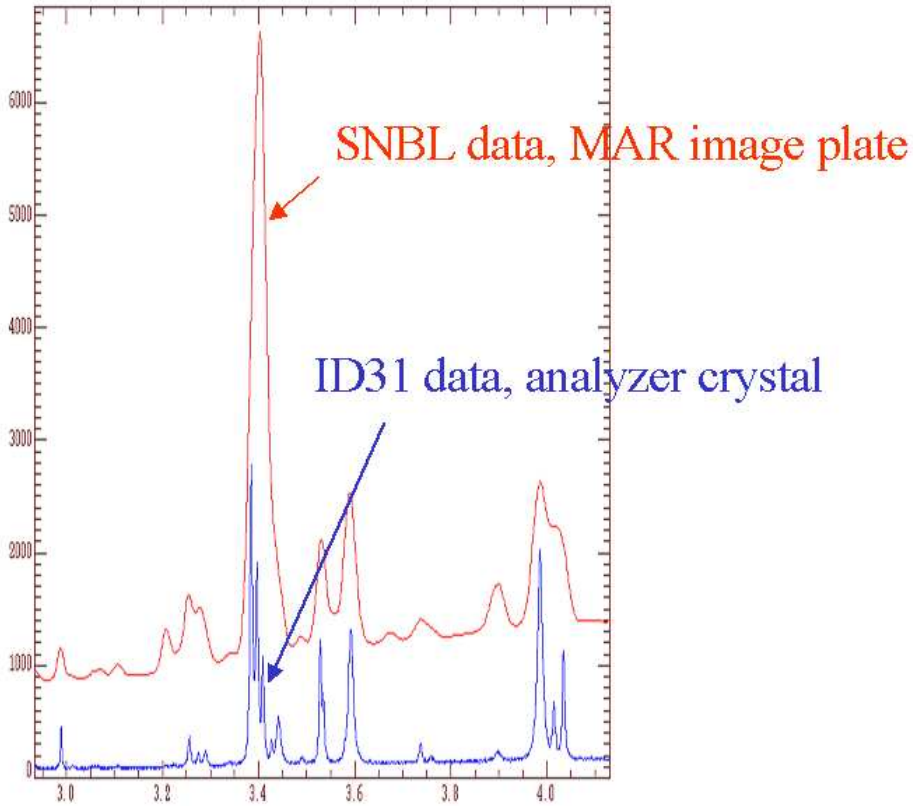


Figure 3-10 Comparison between the difference of resolution of the image plate and the analyser crystal

With FOX we used a global optimisation technique where there is no need to extract the integrated intensities as explained in par 1.2.4 so even with a low resolution diffractogram it was possible to solve the structure.

The final Rietveld refinement was done with the program GSAS[53]. The peaks were described by a pseudo-Voigt peak shape function[54] with asymmetry corrected by the model

of Finger, Cox and Jephcoat[55]. One overall temperature factor was refined, U_{iso}=0.078(1) Å^2. The distances and angles were restrained according to the values found in the microwave spectroscopy work of Creswell[56]. Bond length restraints were weighted 20 with a tolerance of ± 0.001 Å and angles restraints 10 with a tolerance of ± 1°. The final refinement converged to R_{wp} = 10.4%, R_p=7.6% and χ^2=6.0

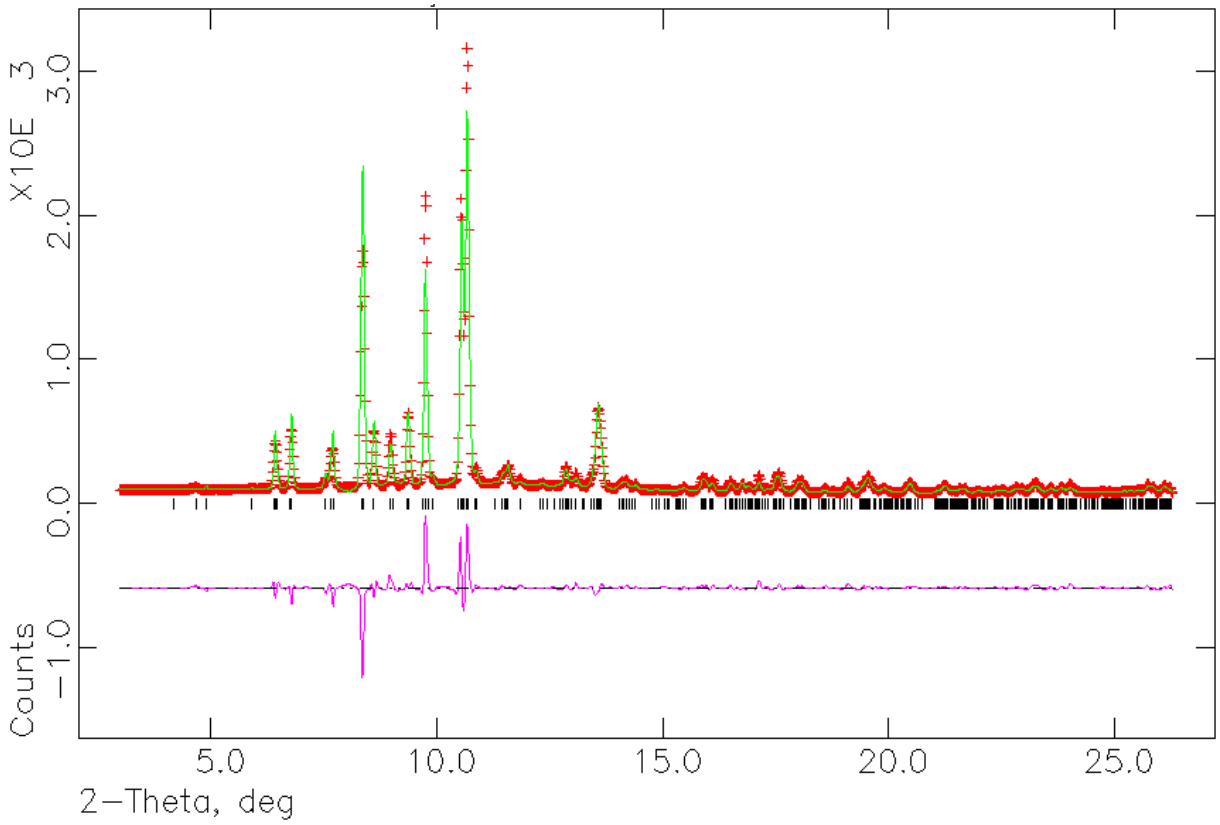


Figure 3-11 Observed (+,red), calculated(-, green) and difference plot of the final Rietveld refinement of phase III, data collected on the SNBL.

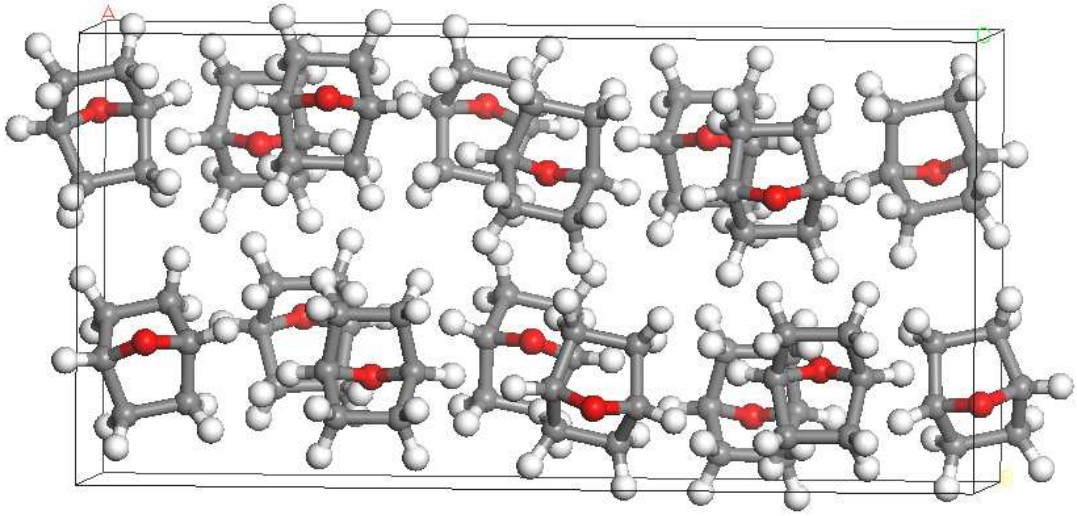
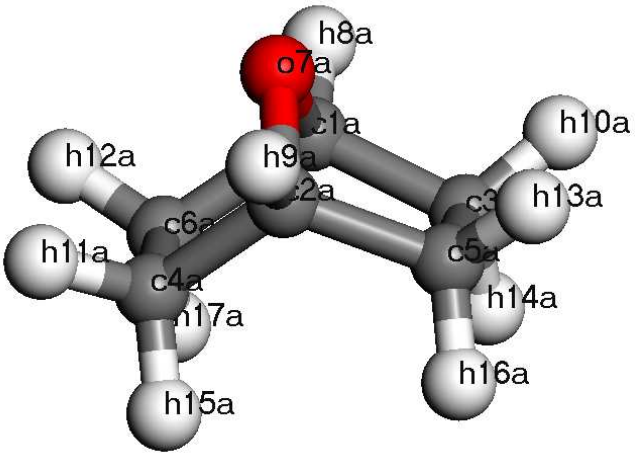


Figure 3-12 Packing of the molecules in the unit cell in the phase III

In Table 3-3 and in Table 3-4 the final distance and angle values are reported



Vector	Length	Vector	Length	Vector	Length	Vector	Length
c1a_c3a	1.544(6)	c1b_c3b	1.540(6)	c1c_c3c	1.535(6)	c1d_c3d	1.534(6)
c1a_c6a	1.540(6)	c1b_c6b	1.542(6)	c1c_c6c	1.550(6)	c1d_c6d	1.546(6)
c1a_o7a	1.445(6)	c1b_o7b	1.447(6)	c1c_o7c	1.446(6)	c1d_o7d	1.449(6)
c1a_h8a	1.080(8)	c1b_h8b	1.080(7)	c1c_h8c	1.080(8)	c1d_h8d	1.080(8)
c2a_c4a	1.537(6)	c2b_c4b	1.546(6)	c2c_c4c	1.545(6)	c2d_c4d	1.546(6)
c2a_c5a	1.544(6)	c2b_c5b	1.541(6)	c2c_c5c	1.538(6)	c2d_c5d	1.543(6)
c2a_o7a	1.454(7)	c2b_o7b	1.447(6)	c2c_o7c	1.455(6)	c2d_o7d	1.453(6)
c2a_h9a	1.080(7)	c2b_h9b	1.080(7)	c2c_h9c	1.080(7)	c2d_h9d	1.080(7)
c3a_c5a	1.548(6)	c3b_c5b	1.547(6)	c3c_c5c	1.545(6)	c3d_c5d	1.545(6)
c3a_h10a	1.081(7)	c3b_h10b	1.080(8)	c3c_h10c	1.079(8)	c3d_h10d	1.079(8)
c3a_h14a	1.080(8)	c3b_h14b	1.079(8)	c3c_h14c	1.079(8)	c3d_h14d	1.080(8)
c4a_c6a	1.544(6)	c4b_c6b	1.544(6)	c4c_c6c	1.548(6)	c4d_c6d	1.545(6)
c4a_h11a	1.080(8)	c4b_h11b	1.080(8)	c4c_h11c	1.080(8)	c4d_h11d	1.080(7)
c4a_h15a	1.080(8)	c4b_h15b	1.079(7)	c4c_h15c	1.080(7)	c4d_h15d	1.080(7)
c5a_h13a	1.080(8)	c5b_h13b	1.081(8)	c5c_h13c	1.080(8)	c5d_h13d	1.081(8)
c5a_h16a	1.080(7)	c5b_h16b	1.080(8)	c5c_h16c	1.080(7)	c5d_h16d	1.080(7)
c6a_h12a	1.080(8)	c6b_h12b	1.079(8)	c6c_h12c	1.080(8)	c6d_h12d	1.080(8)
c6a_h17a	1.080(8)	c6b_h17b	1.080(8)	c6c_h17c	1.080(8)	c6d_h17d	1.079(8)

Table 3-3 Refined bond lengths values

Angle	Degrees	Angle	Degrees	Angle	Degrees	Angle	Degrees
c3a_c1a_c6a	113.9(3)	c3b_c1b_c6b	113.4(5)	c3c_c1c_c6c	114.2(3)	c3d_c1d_c6d	114.8(3)
c3a_c1a_o7a	102.5(4)	c3b_c1b_o7b	101.5(4)	c3c_c1c_o7c	100.6(4)	c3d_c1d_o7d	101.6(4)
c3a_c1a_h8a	113.0(6)	c3b_c1b_h8b	112.8(7)	c3c_c1c_h8c	112.7(7)	c3d_c1d_h8d	106.7(7)
c6a_c1a_o7a	101.1(5)	c6b_c1b_o7b	102.5(4)	c6c_c1c_o7c	102.9(4)	c6d_c1d_o7d	103.6(4)
c6a_c1a_h8a	113.0(6)	c6b_c1b_h8b	112.7(7)	c6c_c1c_h8c	113.1(6)	c6d_c1d_h8d	115.3(6)
o7a_c1a_h8a	112.2(6)	o7b_c1b_h8b	112.4(6)	o7c_c1c_h8c	112.2(6)	o7d_c1d_h8d	114.2(6)
c4a_c2a_c5a	113.7(4)	c4b_c2b_c5b	113.8(3)	c4c_c2c_c5c	114.3(3)	c4d_c2d_c5d	114.5(3)
c4a_c2a_o7a	97.1(5)	c4b_c2b_o7b	103.7(4)	c4c_c2c_o7c	102.7(4)	c4d_c2d_o7d	102.5(4)
c4a_c2a_h9a	110.4(7)	c4b_c2b_h9b	112.6(7)	c4c_c2c_h9c	112.6(7)	c4d_c2d_h9d	112.5(7)
c5a_c2a_h9a	110.7(7)	c5b_c2b_h9b	112.6(7)	c5c_c2c_h9c	112.5(7)	c5d_c2d_h9d	112.5(7)
o7a_c2a_h9a	110.0(7)	o7b_c2b_h9b	111.7(6)	o7c_c2c_h9c	112.3(7)	o7d_c2d_h9d	112.3(7)
c1a_c3a_c5a	100.1(3)	c1b_c3b_c5b	101.5(3)	c1c_c3c_c5c	101.3(3)	c1d_c3d_c5d	99.7(3)
c1a_c3a_h10a	111.4(6)	c1b_c3b_h10b	111.1(6)	c1c_c3c_h10c	110.8(6)	c1d_c3d_h10d	111.1(6)
c1a_c3a_h14a	111.3(6)	c1b_c3b_h14b	110.9(6)	c1c_c3c_h14c	111.3(7)	c1d_c3d_h14d	111.8(7)
c5a_c3a_h10a	111.1(7)	c5b_c3b_h10b	110.8(6)	c5c_c3c_h10c	110.9(7)	c5d_c3d_h10d	110.9(7)
c5a_c3a_h14a	111.2(6)	c5b_c3b_h14b	110.9(6)	c5c_c3c_h14c	110.9(6)	c5d_c3d_h14d	111.3(6)
h10a_c3a_h14a	111.2(8)	h10b_c3b_h14b	111.2(8)	h10c_c3c_h14c	111.2(8)	h10d_c3d_h14d	111.5(8)
c2a_c4a_c6a	100.8(3)	c2b_c4b_c6b	100.6(3)	c2c_c4c_c6c	100.5(3)	c2d_c4d_c6d	100.2(3)
c2a_c4a_h11a	111.1(6)	c2b_c4b_h11b	111.4(6)	c2c_c4c_h11c	111.1(6)	c2d_c4d_h11d	111.2(7)
c2a_c4a_h15a	111.4(7)	c2b_c4b_h15b	111.2(7)	c2c_c4c_h15c	111.5(7)	c2d_c4d_h15d	111.5(7)
c6a_c4a_h11a	111.2(7)	c6b_c4b_h11b	111.1(6)	c6c_c4c_h11c	110.9(6)	c6d_c4d_h11d	110.7(7)
c6a_c4a_h15a	110.9(7)	c6b_c4b_h15b	111.0(6)	c6c_c4c_h15c	111.3(7)	c6d_c4d_h15d	111.2(7)
h11a_c4a_h15a	111.4(7)	h11b_c4b_h15b	111.1(8)	h11c_c4c_h15c	111.2(8)	h11d_c4d_h15d	111.4(8)
c2a_c5a_c3a	100.1(3)	c2b_c5b_c3b	99.8(3)	c2c_c5c_c3c	100.5(3)	c2d_c5d_c3d	101.4(3)
c2a_c5a_h13a	111.3(6)	c2b_c5b_h13b	111.2(6)	c2c_c5c_h13c	111.0(6)	c2d_c5d_h13d	111.1(6)
c2a_c5a_h16a	111.3(7)	c2b_c5b_h16b	111.4(7)	c2c_c5c_h16c	111.4(7)	c2d_c5d_h16d	111.1(7)
c3a_c5a_h13a	110.9(6)	c3b_c5b_h13b	110.9(7)	c3c_c5c_h13c	111.0(6)	c3d_c5d_h13d	110.7(6)
h13a_c5a_h16a	111.1(8)	h13b_c5b_h16b	111.5(8)	h13c_c5c_h16c	111.3(8)	h13d_c5d_h16d	111.3(8)
c1a_c6a_c4a	100.7(2)	c1b_c6b_c4b	100.8(3)	c1c_c6c_c4c	101.1(3)	c1d_c6d_c4d	100.8(3)
c1a_c6a_h12a	111.1(6)	c1b_c6b_h12b	111.0(6)	c1c_c6c_h12c	111.3(6)	c1d_c6d_h12d	111.5(6)
c1a_c6a_h17a	111.3(7)	c1b_c6b_h17b	111.5(7)	c1c_c6c_h17c	111.3(6)	c1d_c6d_h17d	111.2(6)
c4a_c6a_h12a	111.0(6)	c4b_c6b_h12b	110.8(7)	c4c_c6c_h12c	110.8(7)	c4d_c6d_h12d	110.8(6)
c4a_c6a_h17a	110.8(6)	c4b_c6b_h17b	111.0(6)	c4c_c6c_h17c	111.0(6)	c4d_c6d_h17d	110.8(6)
h12a_c6a_h17a	111.3(8)	h12b_c6b_h17b	111.2(8)	h12c_c6c_h17c	111.0(8)	h12d_c6d_h17d	111.2(8)
c1a_o7a_c2a	94.5(4)	c1b_o7b_c2b	94.2(4)	c1c_o7c_c2c	94.4(4)	c1d_o7d_c2d	93.4(4)

Table 3-4 Refined angles values

The forces that hold together the structure are dipole-dipole type interactions.

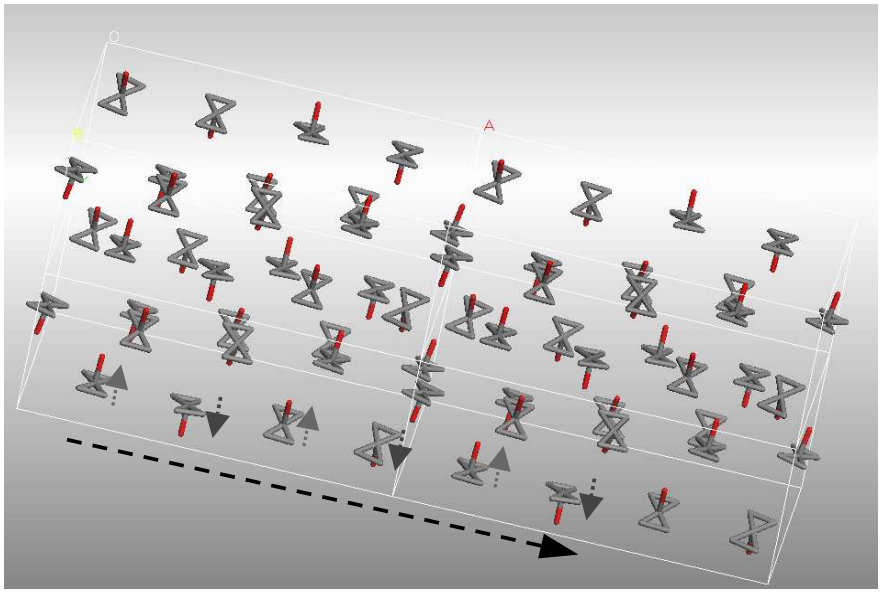


Figure 3-13 Orientation of the dipoles in the phase III

In Figure 3-13 we have sketched the molecules to show more clearly the dipole orientation. The dipoles are alternating up and down in a chain and the crystal structure is formed by a succession of these chains.

	7-oxabicyclo[2.2.1]heptane
molecular formula	$C_6H_{10}O$
molecular weight	98.14
***a*/Å**	21.852(1)
***b*/Å**	10.9699(4)
***c*/Å**	9.3331(4)
β/°	90.453(5)°
space group	$P2_1/a$
V/Å^3	2237.3(1)
Z	16
Z'	4
density/g cm^{-3}	1.165
T/K	185
λ/Å	0.79945
R_{wp}(%)	10.4
R_p(%)	7.6
χ^2	6.0

3.2.1.4 Phase IV

The diffraction pattern recorded on the beamline BM16 at 50 K was indexed from the positions of 20 low-angle peaks using the program KOHL, with a figure of merit M_{20} of 28, in an monoclinic cell:

$$a = 13.6378(1)\ \text{Å} \quad b = 10.2596(1)\ \text{Å} \quad c = 24.0470(2)\ \text{Å} \quad \beta = 102.5217(4)^\circ$$

The cell volume indicates 24 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be C2/c, and hence there are 3 independent molecules in the asymmetric unit.

The structure was solved using the program DASH (simulated annealing). The final Rietveld refinement was done with the program TOPAS. The molecules were treated as a rigid body with values of bond lengths and angles restrained according to the work of Creswell. In the first stage of the refinement only the orientational and positional parameters were allowed to vary. At the end the rigid body was almost completely relaxed, everything was free to vary except the hydrogen atoms, which were kept fixed in the right geometry. Also in this case we use restraints but this time only for the C-C and C-O distances with distances with a tolerance value of ± 0.005 Å and weight 1. The angles were left free to vary. Looking at the data collected on the SNBL we saw that even for this phase, a little preferred orientation is present. Therefore we used preferred orientation correction based on spherical harmonics according to Järvinen[57] in order to improve the fit. We used a 4th-order function and it was enough to correct it. The peaks were described by a PearsonVII[58] (PVII) peak shape function with asymmetry corrected by the simple axial mode present in TOPAS. The temperature factors were refined individually for each molecule reaching these final B values, B_a= 0.29(6) Å^2, B_b= 0.76(7) Å^2, B_c= 0.41(7) Å^2. The final refinement converged to R_{wp} = 8.8%, R_p=6.8% and χ^2=2.6

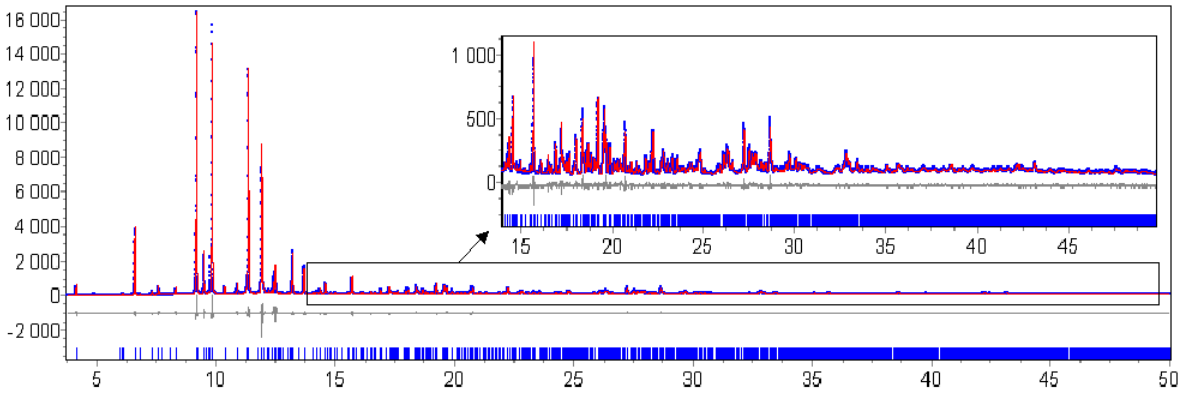


Figure 3-14 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement of phase IV

In the next table we report the final values of bonds and angles, the labels are the same as used in Table 3-3 Table 3-4.

Vector	Length	Vector	Length	Vector	Length
c1a_c3a	1.542(8)	c1b_c3b	1.544(8)	c1c_c3c	1.541(8)
c1a_c6a	1.537(8)	c1b_c6b	1.541(8)	c1c_c6c	1.540(7)
c1a_o7a	1.457(6)	c1b_o7b	1.458(8)	c1c_o7c	1.457(7)
c2a_c4a	1.545(8)	c2b_c4b	1.541(7)	c2c_c4c	1.545(7)
c2a_c5a	1.542(8)	c2b_c5b	1.554(7)	c2c_c5c	1.547(8)
c2a_o7a	1.450(8)	c2b_o7b	1.458(1)	c2c_o7c	1.449(1)
c3a_c5a	1.548(6)	c3b_c5b	1.556(2)	c3c_c5c	1.555(2)
c4a_c6a	1.551(2)	c4b_c6b	1.553(2)	c4c_c6c	1.559(2)

Table 3-5 Refined bond lengths values and

Table 3-6 Refined angles values

Angle	Degrees	Angle	Degrees	Angle	Degrees
c3a_c1a_c6a	110.9(5)	c3b_c1b_c6b	110.6(5)	c3c_c1c_c6c	111.0(5)
c3a_c1a_o7a	102.9(5)	c3b_c1b_o7b	102.8(4)	c3c_c1c_o7c	102.0(4)
c6a_c1a_o7a	101.3(5)	c6b_c1b_o7b	102.3(6)	c6c_c1c_o7c	103.1(6)
c4a_c2a_c5a	106.2(6)	c4b_c2b_c5b	113.6(6)	c4c_c2c_c5c	108.7(7)
c4a_c2a_o7a	103.4(5)	c4b_c2b_o7b	102.6(6)	c4c_c2c_o7c	104.0(6)
c5a_c2a_o7a	104.9(5)	c5b_c2b_o7b	102.4(6)	c5c_c2c_o7c	102.5(5)
c1a_c3a_c5a	101.9(6)	c1b_c3b_c5b	100.5(7)	c1c_c3c_c5c	101.7(6)
c2a_c4a_c6a	100.9(6)	c2b_c4b_c6b	99.1(6)	c2c_c4c_c6c	101.4(6)
c2a_c5a_c3a	100.3(6)	c2b_c5b_c3b	101.2(6)	c2c_c5c_c3c	100.3(6)
c1a_c6a_c4a	101.2(6)	c1b_c6b_c4b	102.7(6)	c1c_c6c_c4c	100.3(6)
c1a_o7a_c2a	95.2(4)	c1b_o7b_c2b	94.4(5)	c1c_o7c_c2c	95.1(5)

In Figure 3-15 we show the final crystal structure:

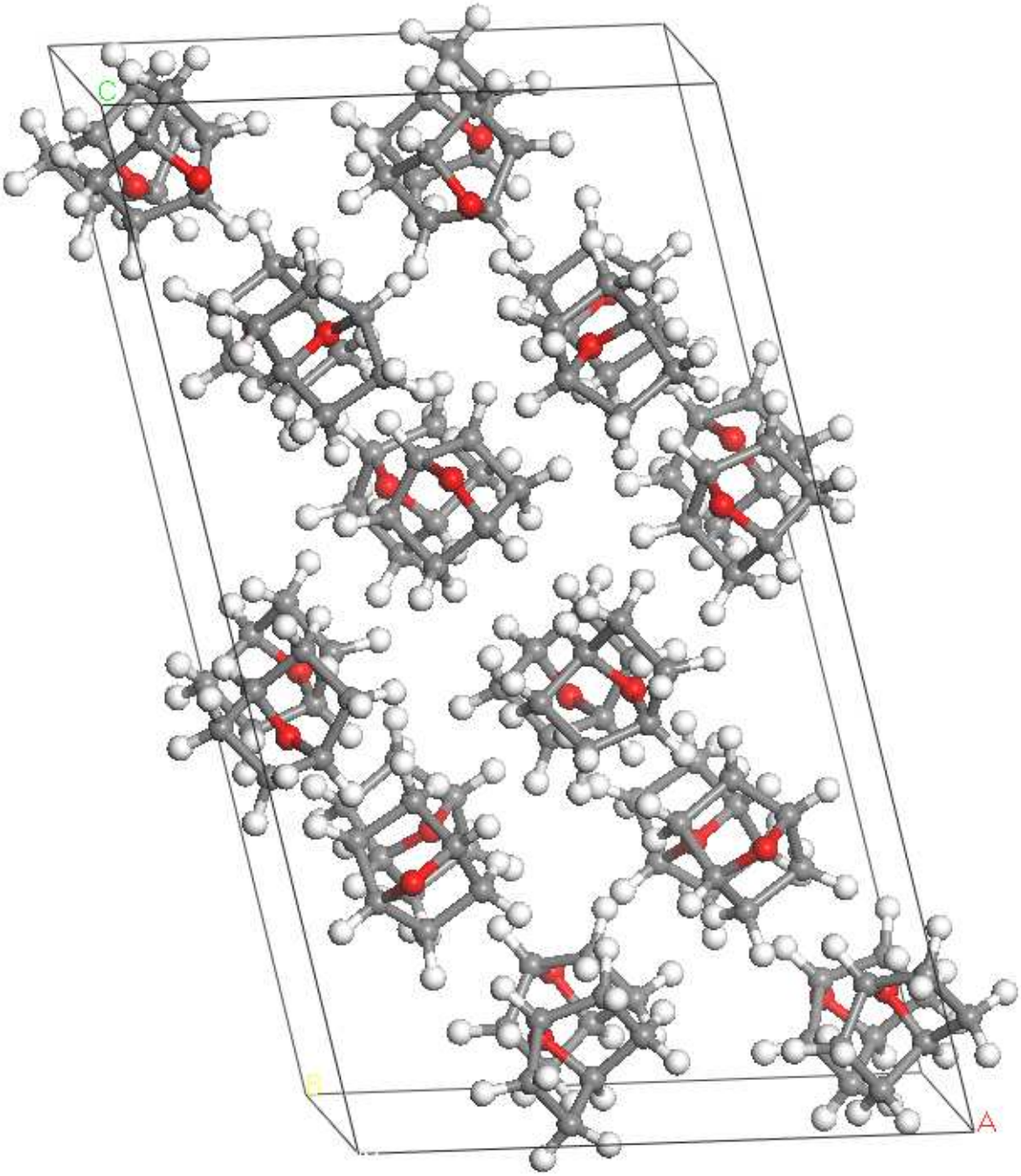


Figure 3-15 Packing of the molecules in the unit cell in the phase IV

The forces that hold together the structure are, as seen before, dipole-dipole type interactions, but the disposition of the dipoles in space is different. Comparing the structure at 185 K, the dipoles are ordered in plane and oriented in opposite directions from one plane to the next.

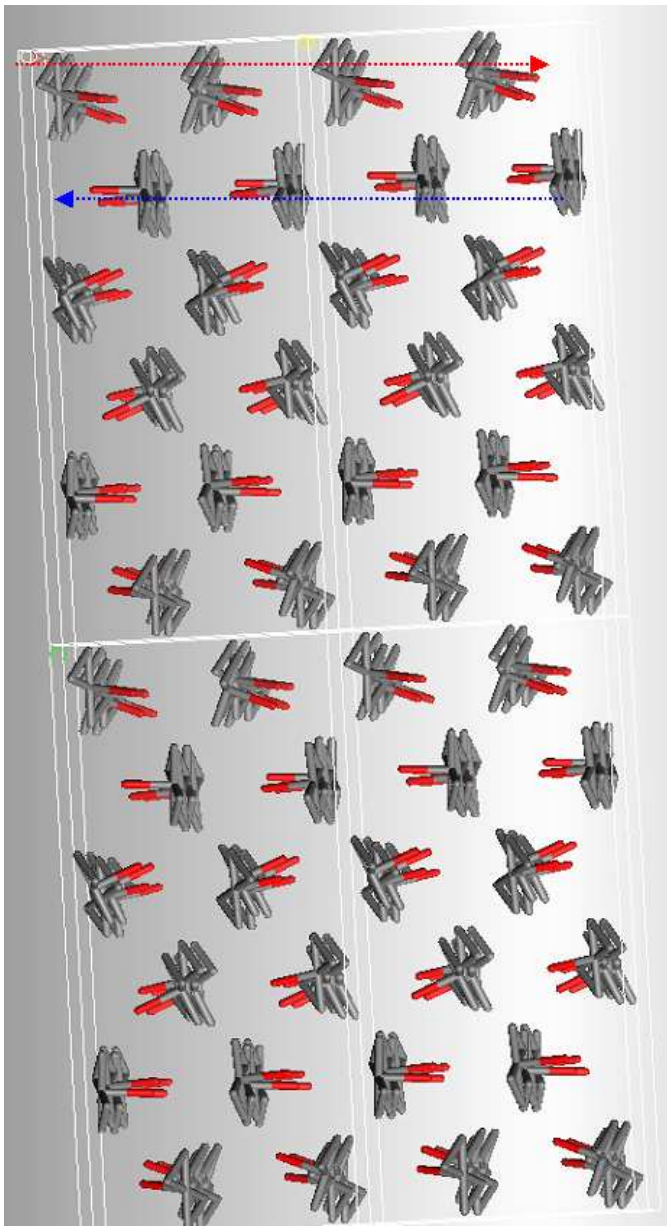
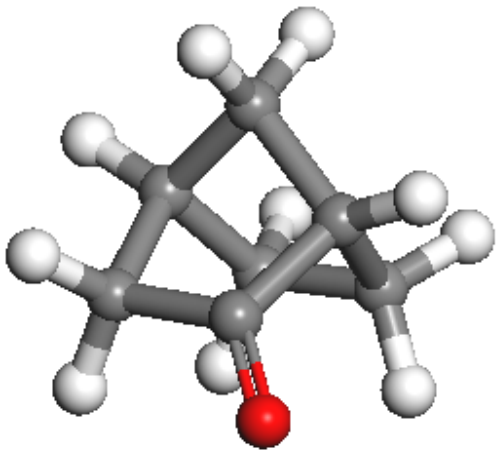


Figure 3-16 Orientation of the dipoles in the phase IV

	7-oxabicyclo[2.2.1]heptane
molecular formula	$C_6H_{10}O$
molecular weight	98.14
***a*/Å**	13.6378(1)
***b*/Å**	10.2596(1)
***c*/Å**	24.0470(2)
β/°	102.5217(4)
space group	*C*2/*c*
V/Å^3	3284.58(6)
Z	24
Z'	3
density/g cm^{-3}	1.191
T/K	50
λ/Å	0.85024
R_{wp}(%)	8.8
R_p(%)	6.4
χ^2	2.6

3.2.2 Norcamphor



The solid-state behaviour of the Norcamphor, $C_7H_{10}O$, was studied by Holderna-Natkaniec et al.[59] using neutron diffraction and inelastic incoherent scattering in the temperature range from 10 K to 290K. AT 290K the sample is solid and forms a disordered plastic crystal, phase I. They observed in cooling at about 140 K a transition in the orientation of the molecules, phase II, and upon heating, the formation of a new phase, phase III, that is stable at low temperature up to 160 K. The inelastic incoherent scattering spectra suggested that phase III was more densely packed than phase II. From the neutron diffraction study they didn't report any cell constant and no attempt to solve the structure was made owing to the poor resolution of the diffractogram.

At room temperature the sample (Aldrich, 98%) is solid, forming a white sticky powder. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from the room-temperature phase to the low-temperature ordered phases was investigated by recording on the beamline BM16 short diffraction patterns of 2 min. each. Cooling at the rate of 4 K min^{-1} we could observe only the transition from the disordered phase I to the ordered phase III, with no evidence for the presence of phase II. Instead, similar to the case of the 7-oxabicyclo[2.2.1]heptane, when warming at a rate of 0.6 K min^{-1} we could observe the formation of phase II. In Table 3-7 we report the temperature at which we observed the transitions.

Warming (0.2 K min^{-1})	T(K)
Ordered phase III	<146
Change	146—153
Ordered phase II	154—164
Disordered phase I	>164

Cooling (4 K min^{-1})	T(K)
Ordered phase III	<134
Disordered phase I	>134

Table 3-7 Temperature at which the phase transition occurs

We can see that also in this case hysteresis is present.

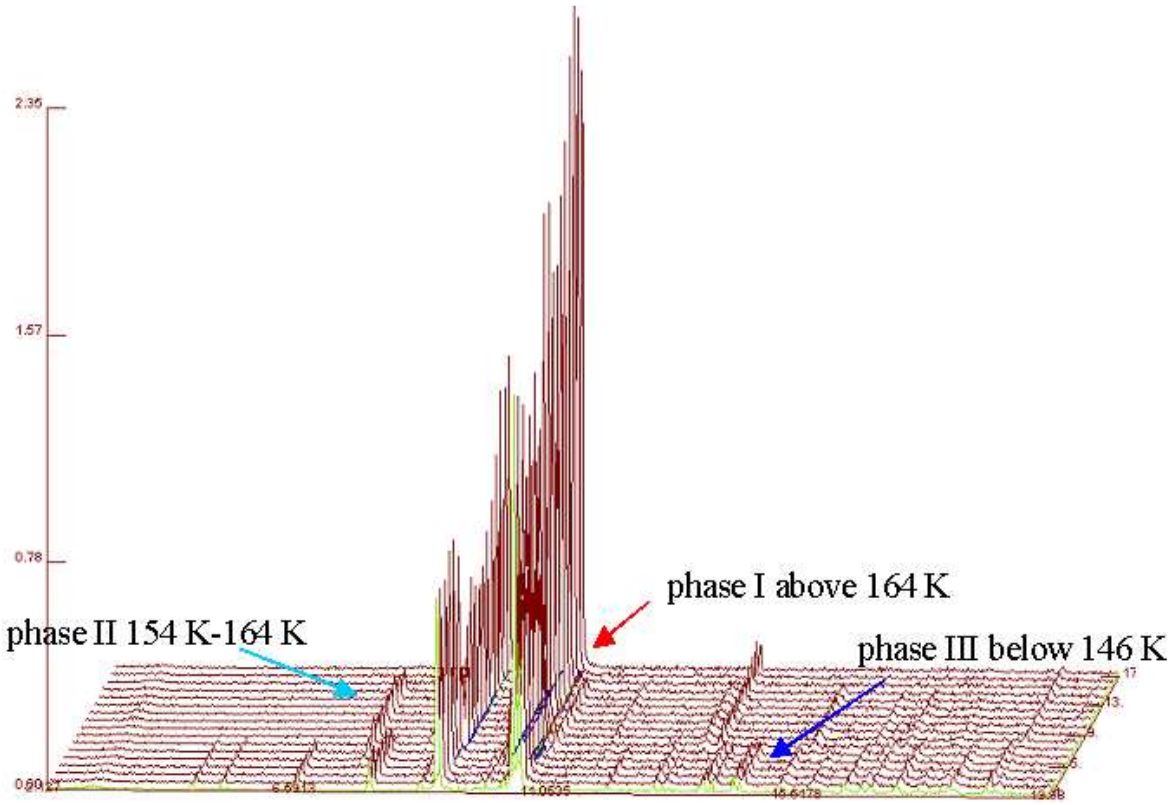


Figure 3-17 Powder diffractograms showing the structural phase transitions

3.2.2.1 Phase I

This phase was manually indexed as hexagonal with a=6.1785(6) Å c=9.7062(9) Å at T=170K. The cell was also checked with the SVD procedures and it found the same cell as we've already determined manually. Looking at the BM16 data the space group can be assign as P63/mmc but a more accurate measurement is required to confirm this hypothesis, as only 5 peaks are visible.

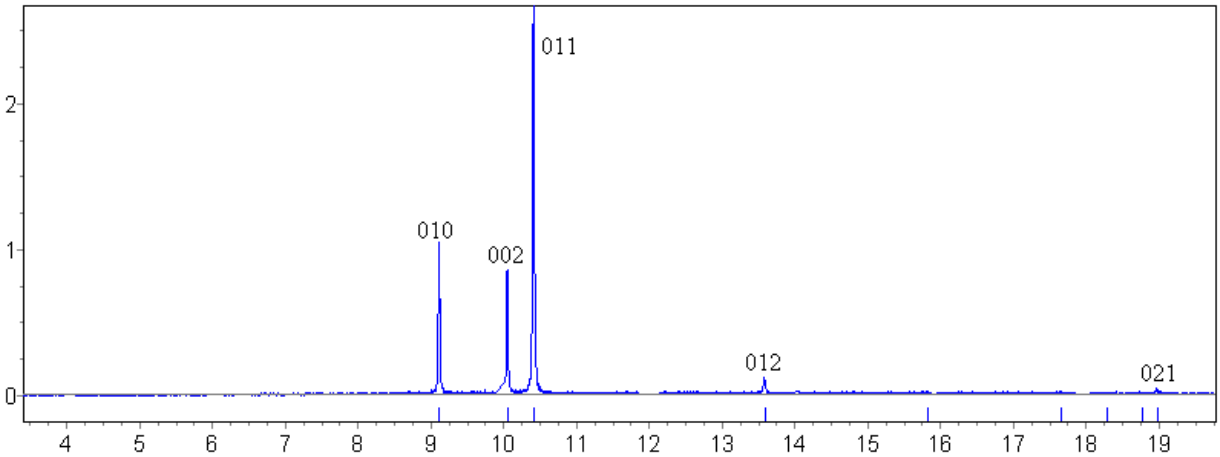


Figure 3-18 Powder diffraction pattern, collected on BM16, of the disordered hexagonal phase I (*λ=0.85024* Å)

3.2.2.2 Phase II

The diffraction pattern recorded on the beamline BM16 at 160 K was indexed from the positions of 20 low-angle peaks using the program ITO, with a figure of merit M_{20} of 31, in a orthorhombic cell:

$$a = 9.1001(3)\ \text{Å}\quad b = 9.3654(3)\ \text{Å}\quad c = 7.3118\ (2)\ \text{Å}$$

The cell volume of 623.19(3) $Å^3$ indicates 4 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $Pna2_1$, and hence there is 1 independent molecule in the asymmetric unit.

Together with the phase II an amount of the hexagonal phase I is still present.

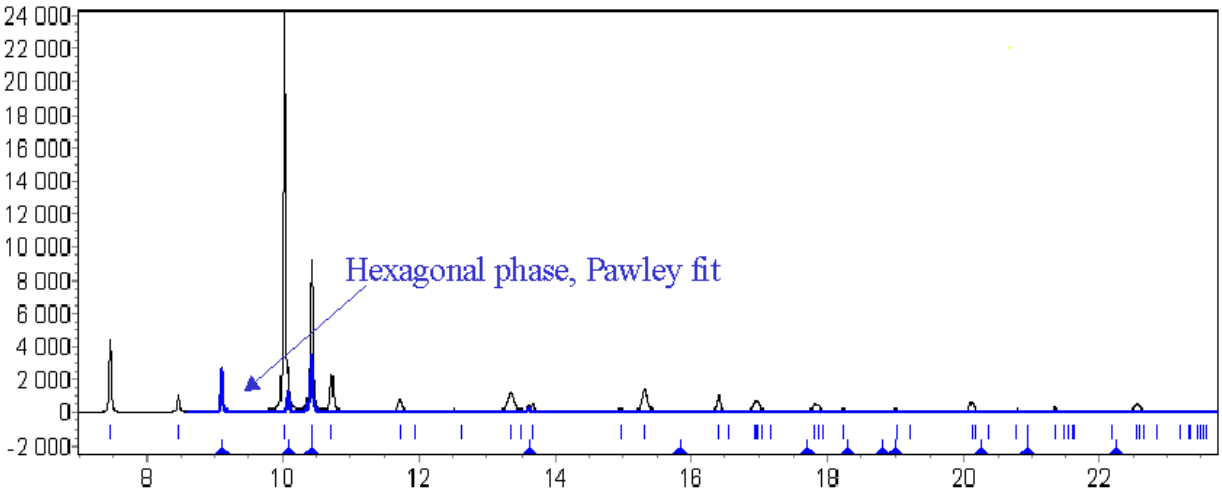


Figure 3-19 Powder diffraction pattern of the phase II, in blue it is showed the contribution of the small amount of phase I still present that is fitted performing a Pawley fit

The peaks intensities of the hexagonal phase I were treated as variables during the structure solution stage performing a Pawley fit. The solution was found with the program Topas, via the simulated annealing procedure when, not only the rotation and translation parameters of the molecule were optimised, but also the intensities of the hexagonal phase. In fact it is not possible to know a priori the right intensities of one of the two phases when there is peak overlap.

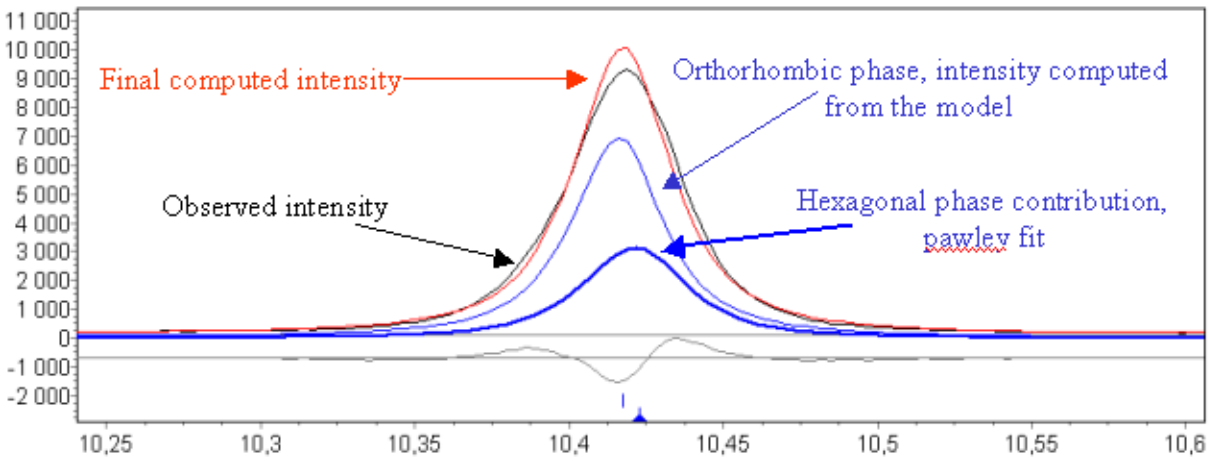


Figure 3-20 Example of how in case of peaks overlap the phase I and phase II contribute respectively

The final Rietveld refinement was also done with the program TOPAS. In the first stage of the refinement only the orientational and positional parameters were allowed to vary. At the end the rigid body was almost completely relaxed; everything was left free to vary except the hydrogen atoms that were kept fix in the right geometry. We use soft restraints for the C-C and C-O distances with a tolerance value of ± 0.01 Å and weight 1. The angles were let free to vary. The peaks were described by a PVII peak shape function with asymmetry corrected by the simple axial mode present in TOPAS. One single temperature factor was refined for the whole molecule reaching the final value, B=6.59(8) Å². The final refinement converged to R_{wp} = 12.5%, R_p=10% and χ^2=3.3

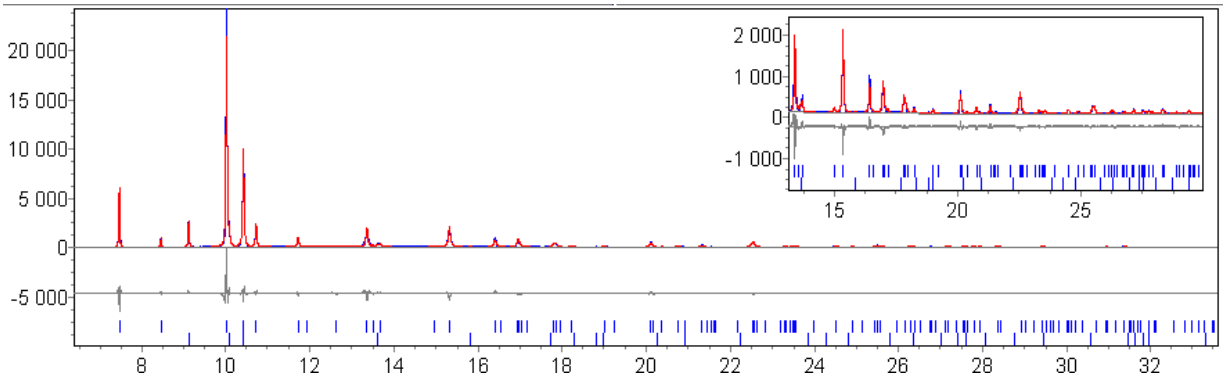


Figure 3-21 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement of phase II

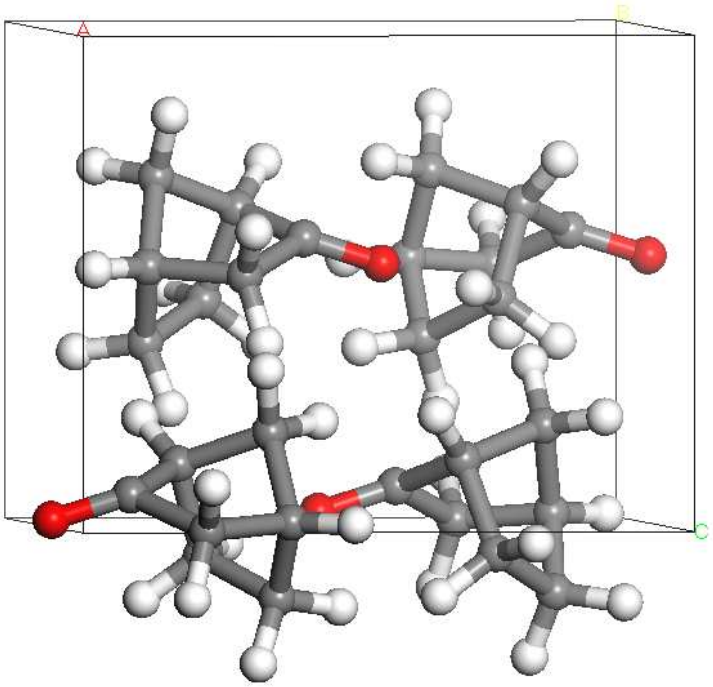


Figure 3-22 Packing of the molecules in the unit cell in the phase II

In Table 3-8 and in Table 3-9 the final distance and angle values are reported

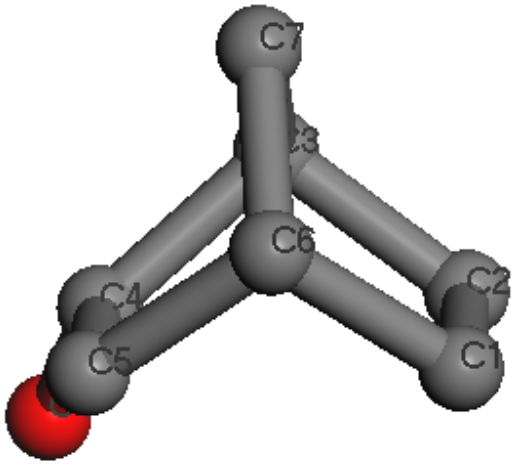


Table 3-8 Refined bond length values

Vector	Length
C1_C2	1.550(9)
C1_C6	1.533(2)
C6_C7	1.527(3)
C7_C3	1.533(2)
C6_C5	1.544(8)
C5_C4	1.540 (1)
C4_O8	1.216(8)
C4_C3	1.526(1)

Angle	Degrees
C6_C1_C2	99.5(1)
C1_C2_C3	106.4(5)
C2_C3_C4	101.7(6)
C2_C3_C7	100.2(4)
C4_C3_C7	101.8(2)
C3_C4_C5	105.1(7)
C3_C4_O8	121.9(9)
C5_C4_O8	126.1(9)
C4_C5_C6	101.1(8)
C5_C6_C7	102.3(1)
C6_C7_C3	93.9(8)

Table 3-9 Refined angles values

Also in this case the forces that determine the crystal packing are dipole-dipole interactions. In the next picture we can see the orientation of the dipole moments in space

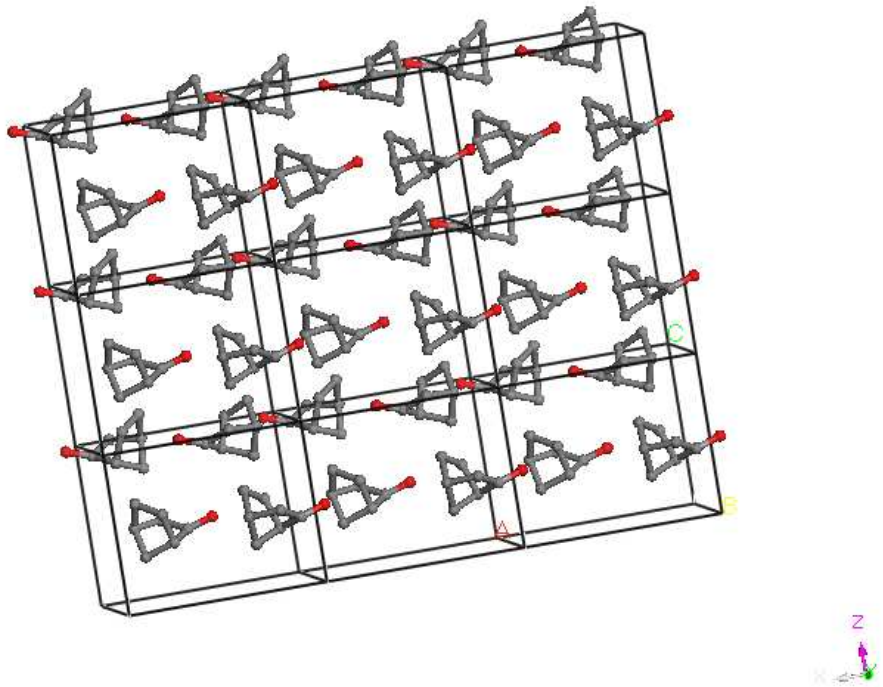


Figure 3-23 Orientation of the dipoles in the phase II

	norcamphor
molecular formula	$C_7H_{10}O$
molecular weight	110.24
***a*/Å**	9.1001(3)
***b*/Å**	9.3654(3)
***c*/Å**	7.3118 (2)
space group	*Pna2₁*
V/Å³	623.19(3)
Z	4
Z'	1
density/g cm⁻³	1.174
T/K	160
λ/Å	0.85024
R_{wp}(%)	12
R_p(%)	10
χ^2	3.3

3.2.2.3 Phase III

The diffraction pattern recorded on the beamline BM16 at 35 K was indexed from the positions of 20 low-angle peaks using the program KOHL, with a figure of merit M_{20} of 54, in a monoclinic cell:

$$a = 18.464(1)\ \text{Å}\quad b = 12.2889(7)\ \text{Å}\quad c = 10.6852(7)\ \text{Å}\quad \beta = 90.993(3)^\circ$$

The cell volume of 2424.1(3) Å^3 indicates 16 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be C2/c, and hence there are 2 independent molecules in the asymmetric unit. In order to solve the structure we used DASH, FOX and TOPAS. None of these attempts was able to find a good solution. We then tried to lower the symmetry to see if we had chosen a wrong space group, then we performed again different computations with space group Cc, with 4 independents molecules. The solution improved but was still not good especially in the high 2θ region, (low d-spacing better spatial resolution). The center of mass of the molecules is almost right but the orientation is still not good. We think that this is due to strong texture in the sample that affects the real value of the observed intensities. The fact that the solution and the final refinement are better in the space group Cc than C2/c could be due to the fact that there are more variables which can take into account in some way the presence of the preferred orientation. We will show here the best fit that we could obtain with TOPAS in the space group Cc. The 4 independent molecules were treated as rigid bodies with bond distances and angle values fixed for the final refinement of the orthorhombic structure at 160 K. In order to compute the preferred orientation effects we used the preferred orientation correction based on spherical harmonics according to Järvinen, via a fourth-order function. One single temperature factor was refined for the whole molecule reaching the final value, B=5.9(1) Å^2. The final refinement converged to R_{wp} = 13.1%, R_p=9.4% and χ^2=4.6

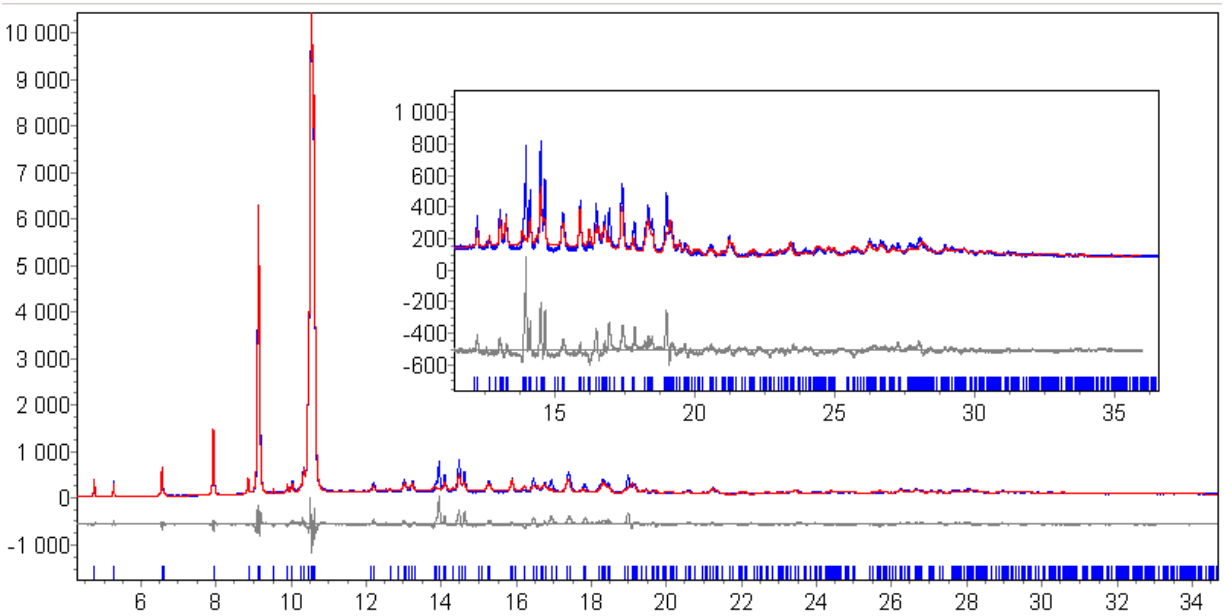


Figure 3-24 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement of phase III

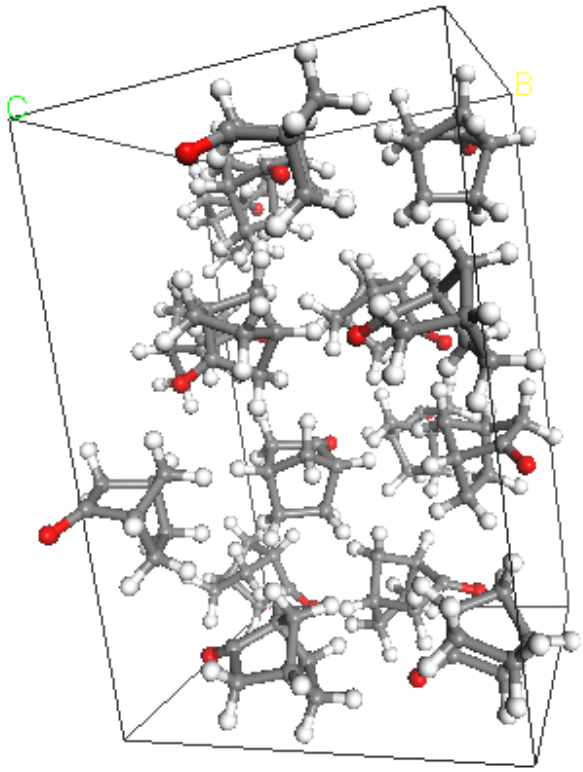


Figure 3-25 The final structure of the phase III of the norcamphor, only the position of the center of mass of the molecules is meaningful

We also performed an experiment on beamline ID11 were we measured this phase at 120K using a CCD area detector. We could then check the poor quality of the powder affected by granularity as we can see in Figure 3-26. Even with this data it was not possible to solve completely the structure.

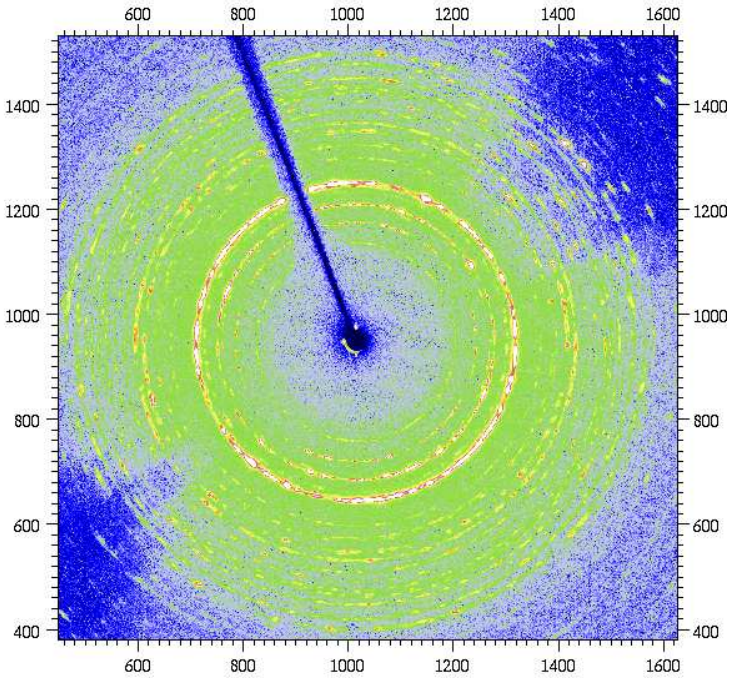
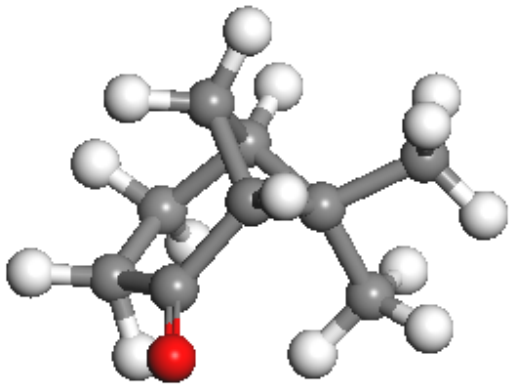


Figure 3-26 Data collected on the beamline ID11 using a CCD detector

	norcamphor
molecular formula	$C_7H_{10}O$
molecular weight	110.24
***a*/Å**	18.464(1))
***b*/Å**	12.2889(7)
***c*/Å**	10.6852(7)
β/°	90.993(3)
space group	*Cc(?)*
V/Å^3	2424.1(3)
Z	16
Z'	4
density/g cm^{-3}	1.208
T/K	35
λ/Å	0.85024
R_{wp}(%)	13.1
R_p(%)	9.4
χ^2	4.6

3.2.3 (1R)-(+)-nopinone



The solid-state behaviour of the (1R)-(+)-nopinone, $C_9H_{14}O$, (Aldrich, 98%) was completely unknown. At room temperature the sample is liquid. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from room temperature to the low-temperature phases was investigated by recording short diffraction patterns of 1 min. each on the beamline ID31. Cooling at the rate of 80 K hour^{-1} we could observe the transition from the liquid to the ordered phase II occurring at T = 250 K. No evidence was seen for any plastic phase. As phase II was affected by preferred orientation effect, it was not possible to solve the structure with the ID31 data so further measurements were made on the SNBL and ID11. From the measurement on the SNBL, where we checked again the evolution of the sample from RT to low temperature, we could detect the presence of the plastic phase I.

Cooling (360 K h^{-1})SNBL	T(K)
Liquid	>260
Disordered phase I	260—240
Ordered phase II	<240

Table 3-10 Temperature at which the phase transitions occur

3.2.3.1 Phase I

Observing the images collected on the SNBL we could observe that at the temperature of 260 K a few single crystals started to grow inside the liquid.

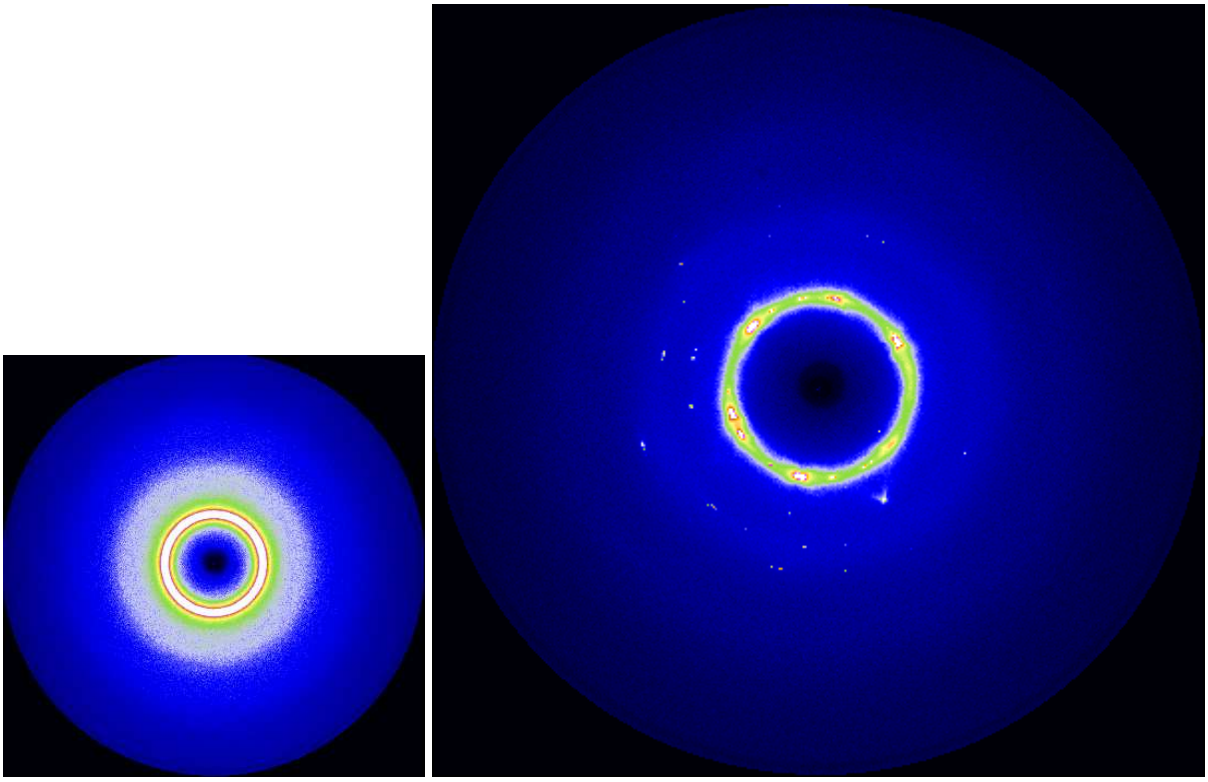


Figure 3-27 Images collected on the SNBL, on the left the liquid phase, on the right is possible to see some spots due to the formation in the liquid of some crystals of the disordered cubic phase.

Looking at the position of the diffraction spots it was possible to index the cell as cubic with a=7.6439 Å at T=255 K.

3.2.3.2 Phase II

The diffraction pattern recorded on ID31 at 120 K was indexed from the positions of 20 low-angle peaks using the program DICVOL, with a figure of merit M_{20} of 42, in an orthorhombic cell:

$$a = 17.65545(6)\ \text{Å}\quad b = 6.96785(3)\ \text{Å}\quad c = 6.52987(2)\ \text{Å}$$

The cell volume of 803.246(5) Å^3 indicates 4 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_12_12_1$, and hence there is 1 independent molecule in the asymmetric unit. In order to solve the structure we used DASH, FOX and TOPAS. None of these attempts was able to find a good solution. We measured then the sample on the SNBL and we saw that it presented big preferred orientation effects.

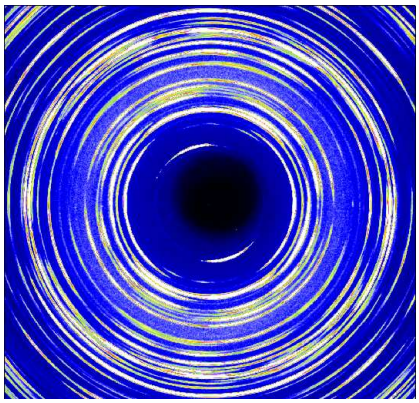


Figure 3-28 Images collected on the SNBL of the phase II, preferred orientations are present in the powder

This image was collected rotating the sample around its axis by 10°. We collected then an image doing a 60° rotation in order to try to average the intensities better but we discovered that even with such a dataset it was not possible to solve the structure.

We performed then another measurement on ID11 with a CCD area detector. In this case we collected sequentially 6 images with a 60° rotation of the capillary around its axis and 3 of 10° rotating the capillary around an axis inclined at 54° with respect to the capillary axis. All the images were finally added together

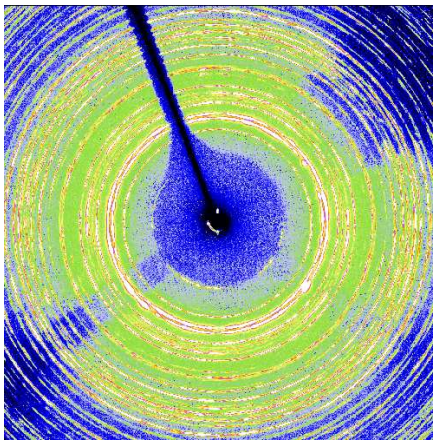


Figure 3-29 Images collected on ID11 rotating the capillary in different directions in order to try to average better the preferred orientation effect

The structure was solved with this last dataset and refined using the program TOPAS. In the first stage of the refinement only the orientational and positional parameters were allowed to vary. At the end the rigid body was almost completely relaxed; everything was free to vary except the hydrogen atoms, which were kept fixed in the right geometry. We use restraints but this time only for the C-C and C-O distances with distances with a tolerance value of ± 0.01 Å and weight 1. The angles were left free to vary. The peaks were described by a PVII peak shape function. One single temperature factors was refined for the whole molecule reaching the final value of B=1.6(1) Å^2. The final refinement converged to R_{wp} = 3.5%, R_p=2.5% and χ^2=1.3

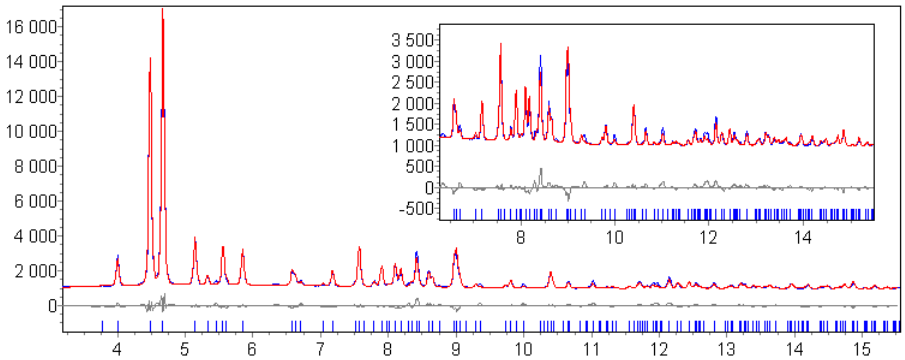


Figure 3-30 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement of phase III

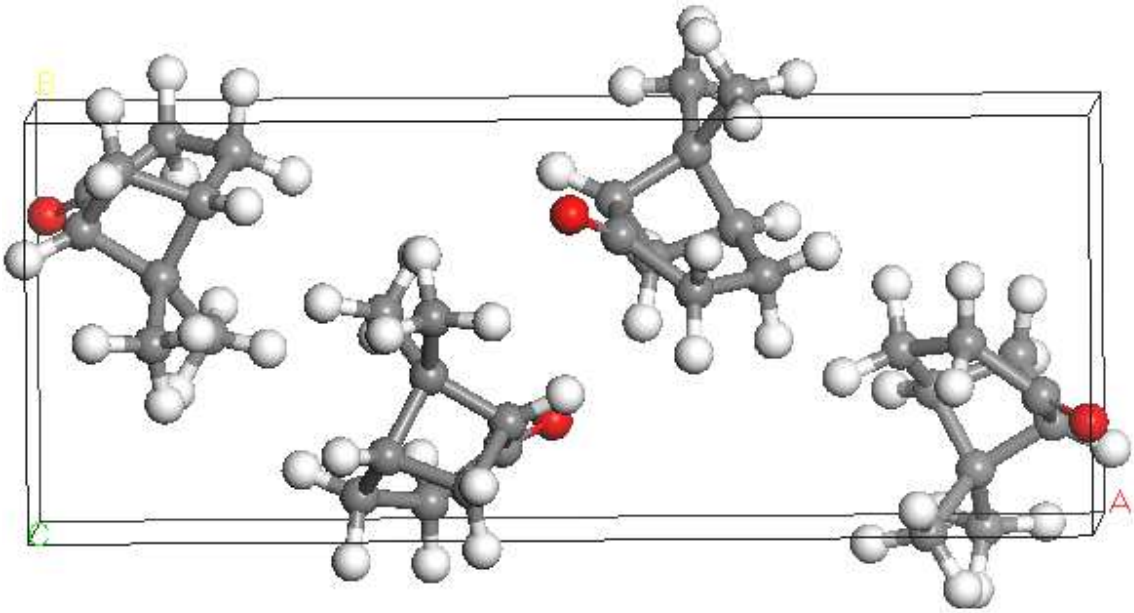
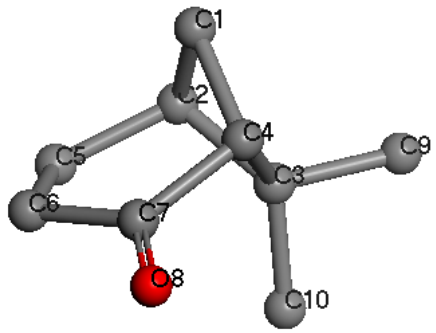


Figure 3-31 Packing of the molecules in the unit cell in the phase II

In Table 3-11 the final distance and angle values are reported



Vector	Length
C7_O8	1.226(15)
C7_C6	1.549(96)
C7_C4	1.554(13)
C6_C5	1.534(16)
C5_C2	1.545(17)
C2_C3	1.532(17)
C2_C1	1.543(15)
C3_C4	1.539(21)
C3_C9	1.533(18)
C3_C10	1.529(19)
C4_C1	1.547(41)

Angle	Degrees
C4_C7_C6	113(2)
C7_C6_C5	114(3)
C6_C5_C2	112(1)
C2_C1_C4	88(1)
C4_C3_C2	89(1)
C10_C3_C9	91(2)
C1_C2_C3	85(1)
C7_C4_C3	104(2)
C4_C7_O8	118(2)
C6_C7_O8	127(3)
C1_C4_C3	85(2)

Table 3-11 Refined bond and angle values

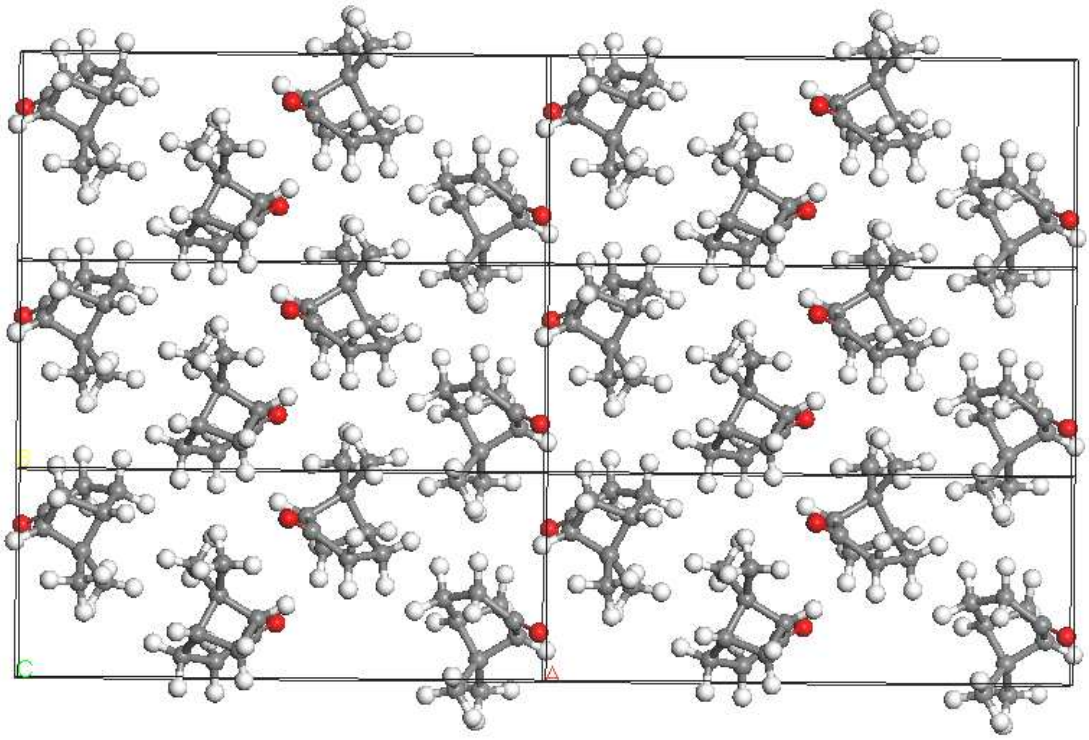
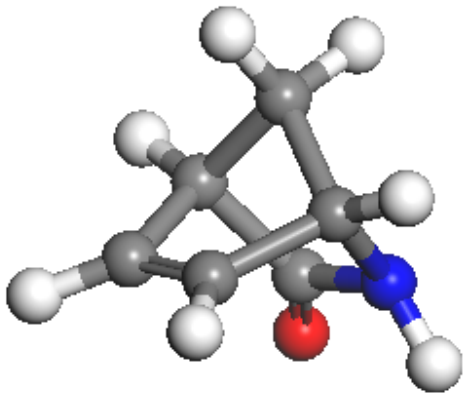


Figure 3-32 Orientation of the dipoles in the phase II

The forces that hold together the structure are dipole-dipole interaction but a big contribution is also coming from repulsion forces between the molecules

	(1R)-(+)-nopinone
molecular formula	$C_9H_{14}O$
molecular weight	138.17
***a*/Å**	17.65545(6)
***b*/Å**	6.96785(3)
***c*/Å**	6.52987(2)
space group	$P2_12_12_1$
V/Å^3	803.246(5)
Z	4
Z'	1
density/g cm^{-3}	1.142
T/K	120
λ/Å	0.42780
R_{wp}(%)	3.5
R_p(%)	2.5
χ^2	1.3

3.2.4 (±)2-Azabicyclo[2.2.1]hept-5-en-3-one



The solid-state behaviour of the (±)2-Azabicyclo[2.2.1]hept-5-en-3-one, C_6H_7NO, (Aldrich,98%) was completely unknown. At room temperature the sample is solid, forming a normal dry powder not sticky as other plastic crystals. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample from room to low temperature was investigated by recording on beamline ID31 short diffraction patterns of 1 min each. No phase transition was observed.

The diffraction pattern recorded on ID31 at 80K was indexed from the positions of 20 low-angle peaks using the program DICVOL, with a figure of merit M_{20} of 26, in an orthorhombic cell:

$$a = 6.2179(1)\ \text{Å}\quad b = 14.3605(3)\ \text{Å}\quad c = 6.0628(1)\ \text{Å}$$

The cell volume of 541.37(2) Å^3 indicates 4 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_12_12_1$, and hence there is 1 independent molecule in the asymmetric unit. The structure was solved and refined using the program TOPAS. In the first stage of the refinement only the orientational and positional parameters were allowed to vary. At the end the rigid body was almost completely relaxed, everything was free to vary except the hydrogen atoms, which were kept fixed in the right geometry. We use restraints but this time only for the C-C and C-O distances with distances with a tolerance value of ± 0.01 Å and weight 1. The angles were left free to vary. The peaks were described by a PVII peak shape function. One single temperature factors was refined

for the whole molecule reaching the final value of B=1.56(5) Å². The final refinement converged to R_{wp} = 10.1%, R_p=8.9% and χ^2=1.6

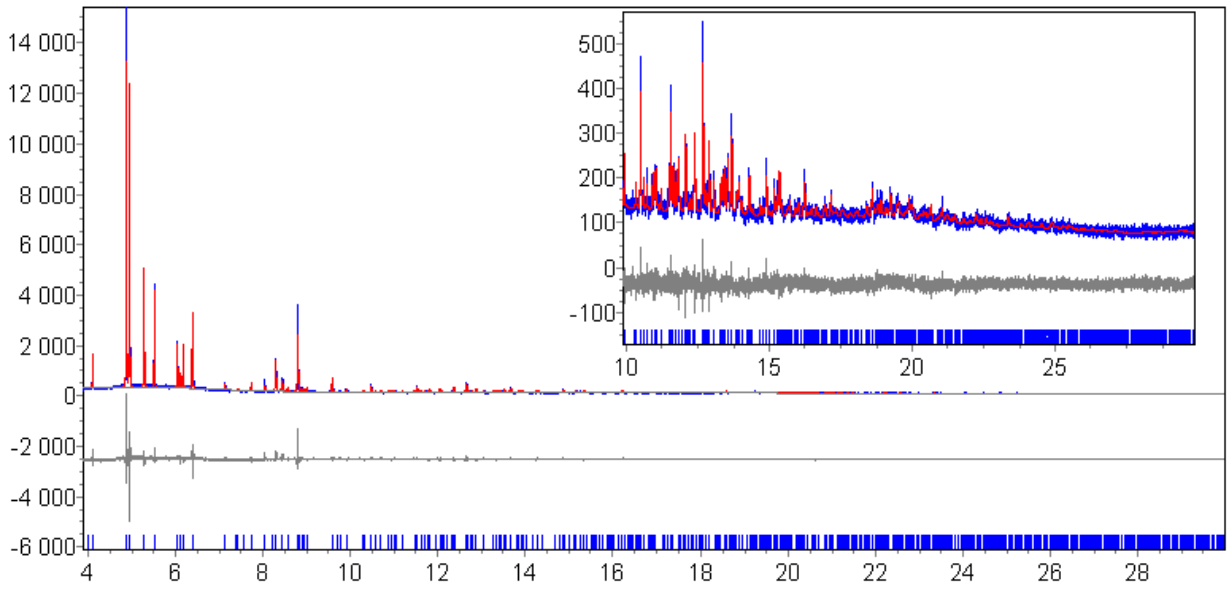


Figure 3-33 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement

In Table 3-12 Table 3-13 the final distance and angle values are reported

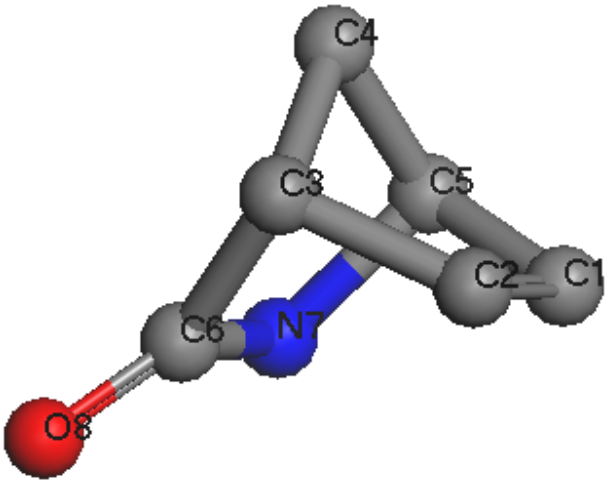


Table 3-12 Refined bond lengths values

Vector	Length
C2_C1	1.325(8)
C2_C3	1.538(10)
C4_C3	1.534(8)
C1_C5	1.524(8)
C6_C3	1.540(9)
N7_C6	1.372(8)
O8_C6	1.202(7)
N7_C5	1.483(4)
C5_C4	1.535(4)

Angle	Degrees
C2_C3_C6	101.6(5)
C2_C3_C4	99.2(6)
C3_C6_N7	105.4(7)
C3_C6_O8	129.2(4)
C6_N7_C5	107.0(1)
N7_C5_C1	102.7(1)
C1_C5_C4	100.2(1)
C1_C2_C3	107.4(8)
C3_C4_C5	92.3(1)
C6_C3_C4	100.5(1)
N7_C5_C4	102.2(1)

Table 3-13 Refined angles values

The forces that hold together the structure are hydrogen bonding type interaction between the oxygen and the nitrogen atoms. This is the only sample that shows this kind of interaction and in fact is the only one that doesn't present a plastic phase.

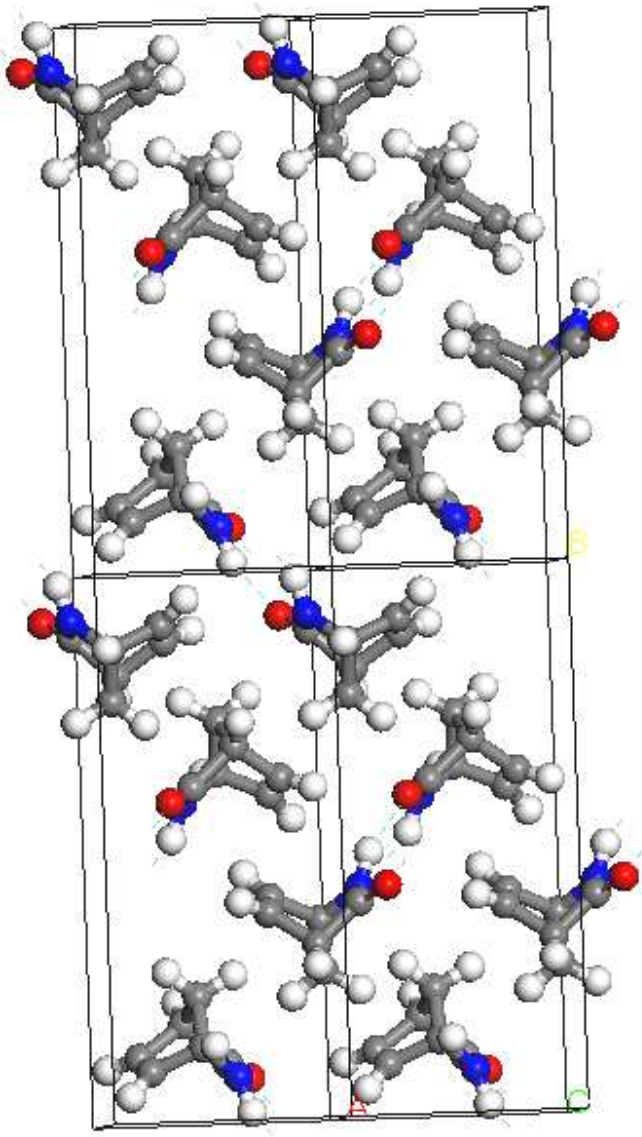
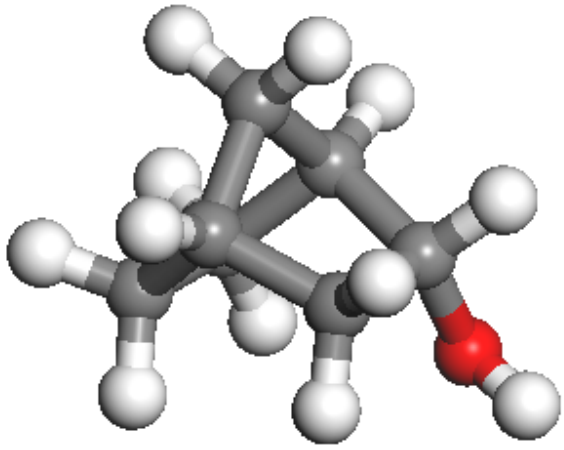


Figure 3-34 Packing of the molecules, the blue dashed line shows the hydrogen bond between the molecules.

Because of the hydrogen bonding the density is higher than the other compounds that form a plastic phase.

	(±)2-Azabicyclo[2.2.1]hept-5-en-3-one
molecular formula	C_6H_7NO
molecular weight	109.13
***a*/Å**	6.2179(1)
***b*/Å**	14.3605(3)
***c*/Å**	6.0628(1)
space group	$P2_12_12_1$
V/Å^3	541.37(2)
Z	4
Z'	1
density/g cm^{-3}	1.338
T/K	80
λ/Å	0.42780
R_{wp}(%)	3.5
R_p(%)	2.5
χ^2	1.6

3.2.5 (+)-endo-2-norborneol



The solid-state behaviour of the (+)-endo-2-norborneol, $C_7H_{12}O$, was completely unknown. At room temperature the sample is solid, forming a white sticky powder. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from the room temperature phase to the low

temperature ordered phases was investigated by recording on beamline ID31 short diffraction patterns of 1 min each. Cooling at the rate of 80 K hour^{-1} we could observe the transition from the disordered phase I to the ordered phase II occurring at T = 240 K.

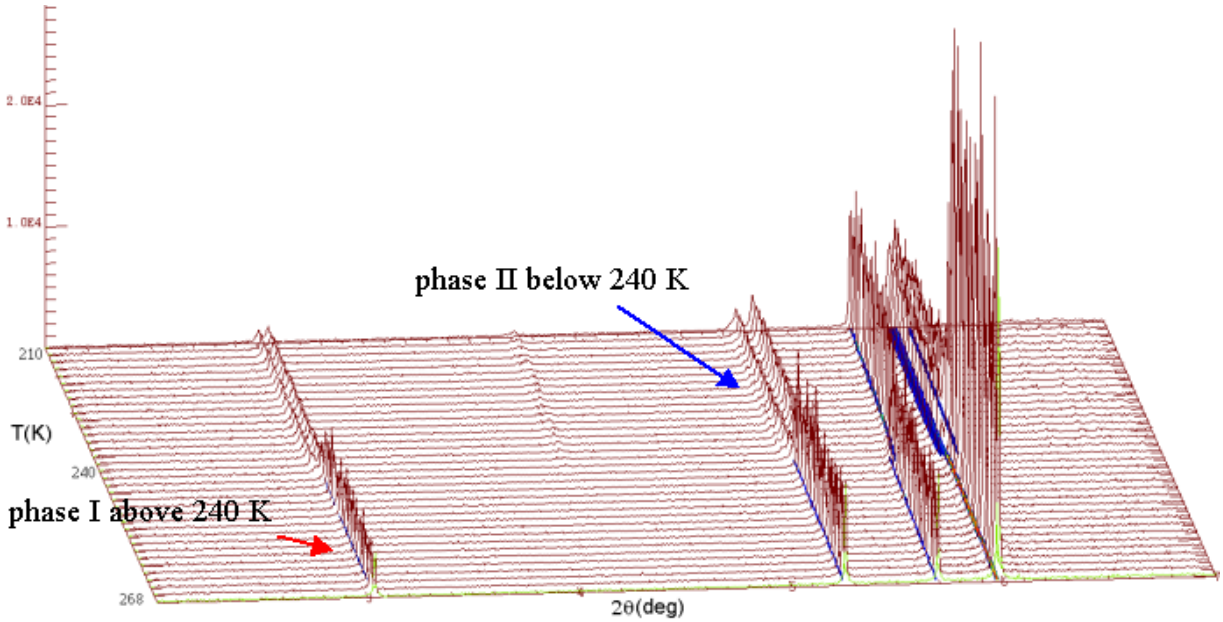


Figure 3-35 Powder diffractograms showing the phase transition

3.2.5.1 Phase I

This phase is hexagonal with a=10.9367(5) Å c=10.1280(5) Å at room temperature, this unit cell was found using the program SVD. The space group is P63.

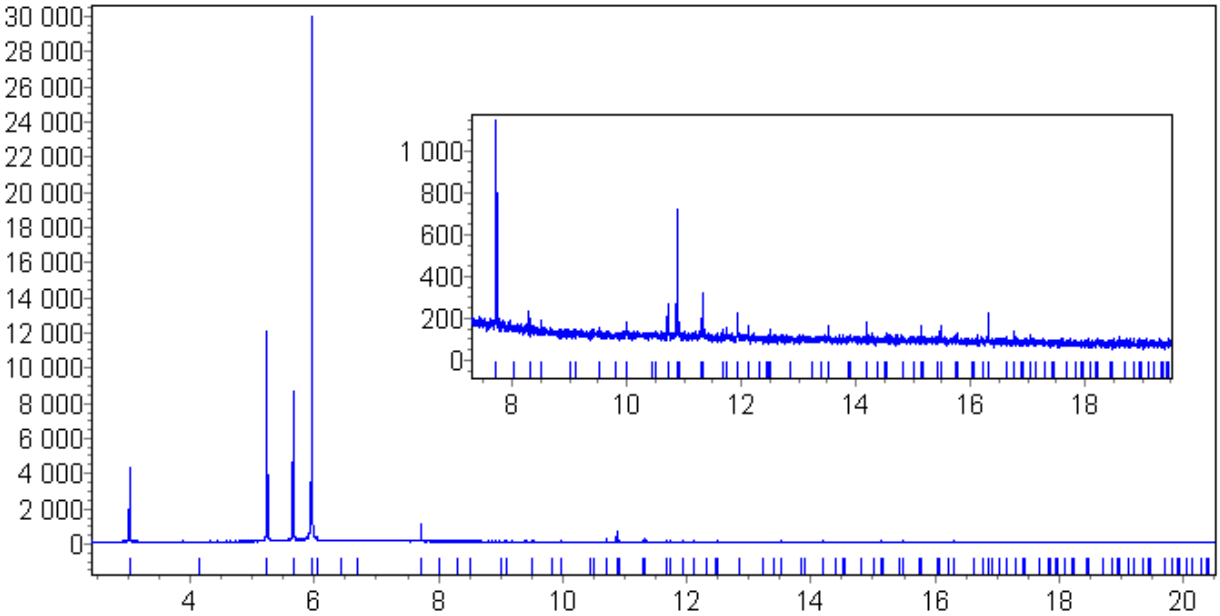


Figure 3-36 Powder diffraction pattern of the disordered hexagonal phase I (λ=*0.50027* Å)

In this case the number of observed peaks is much higher then usually for this kind of compounds. We tried then to solve the structure to see if it was possible to describe the plastic disordered phase. We didn't succeed in the solution so we measured this sample also on the SNBL. As can be seen in Figure 3-37 the sample is affected by strong texture and even averaging over all the diffraction circles was not enough to have the right set of intensities

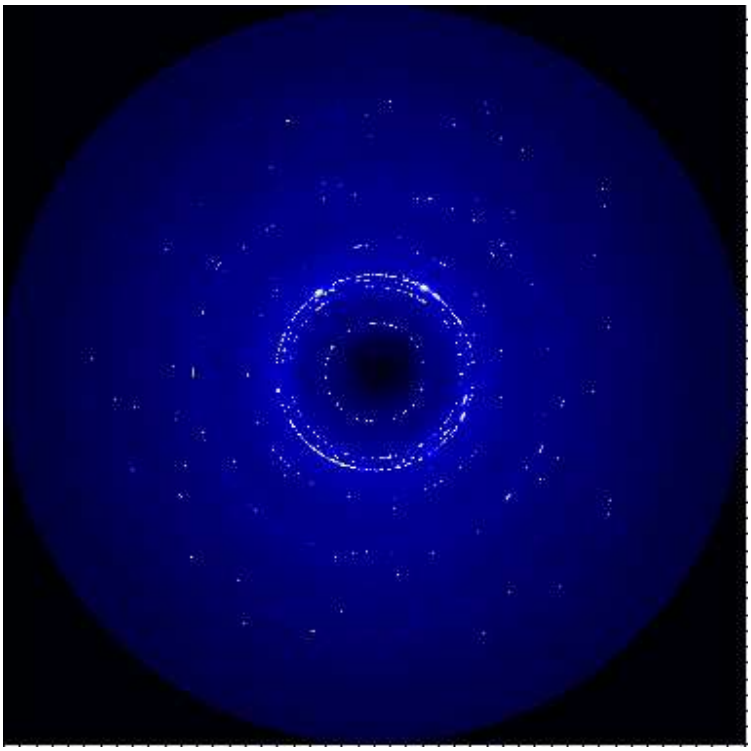


Figure 3-37 Image collected on the SNBL that show the poor quality of the powder

3.2.5.2 Phase II

The diffraction pattern recorded on the beamline ID31 at 80 K was indexed from the positions of 20 low-angle peaks using the program KHOL, with a figure of merit M_{20} of 76, in a monoclinic cell:

$$a = 10.7403(2)\ \text{Å}\quad b = 9.7565(1)\ \text{Å}\quad c = 10.6011(2)\ \text{Å}\quad \beta = 119.301(1)^\circ$$

The cell volume of 968.75(5) Å^3 indicates 6 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_1$, and hence there are 3 independent molecules in the asymmetric unit. In order to solve the structure we used DASH, FOX and TOPAS. None of these attempts was able to find a good solution. As for the phase I the sample is still affected by strong texture effects as can be see form these images colleted at the SNBL

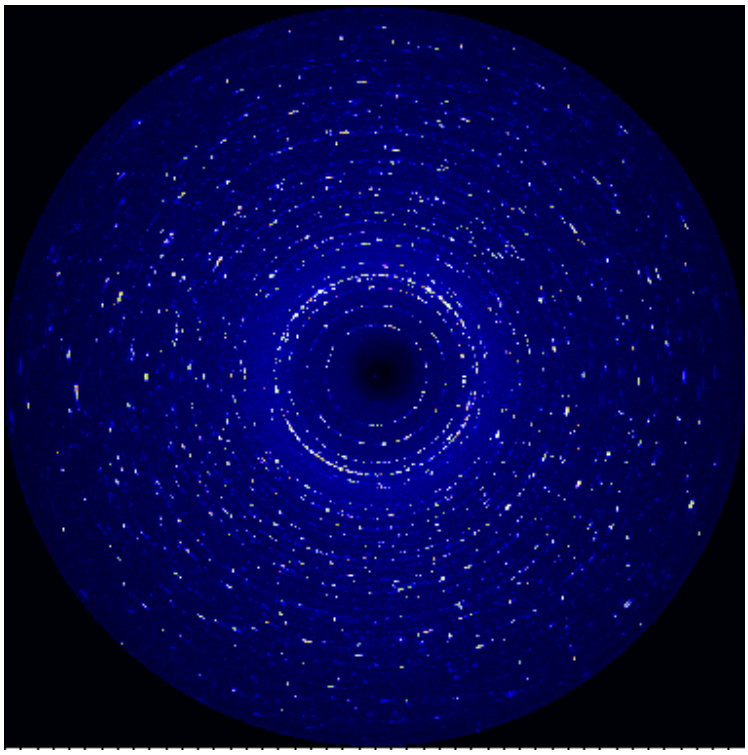


Figure 3-38 Phase II measured on the SNBL

Then we performed another measurement on ID11 using the same data collection strategy as done for the (1R)-(+)-nopinone. But even with this data it was not possible to solve the structure

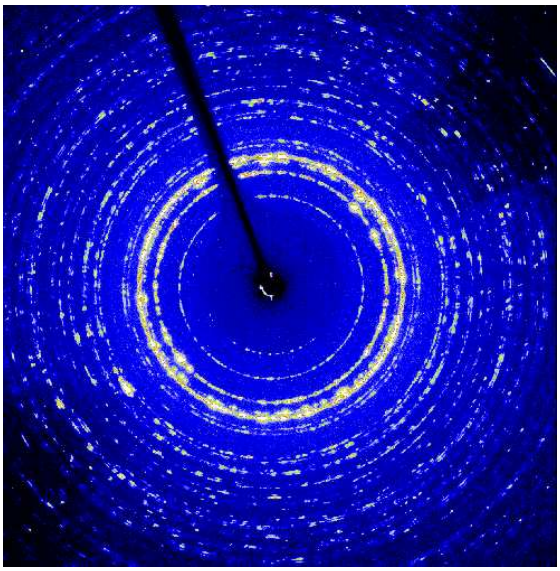


Figure 3-39 Phase II measured on ID11

in Figure 3-40 we show how the poor quality of the powder can dramatically affect the intensity values, this is why is not possible to solve the structure

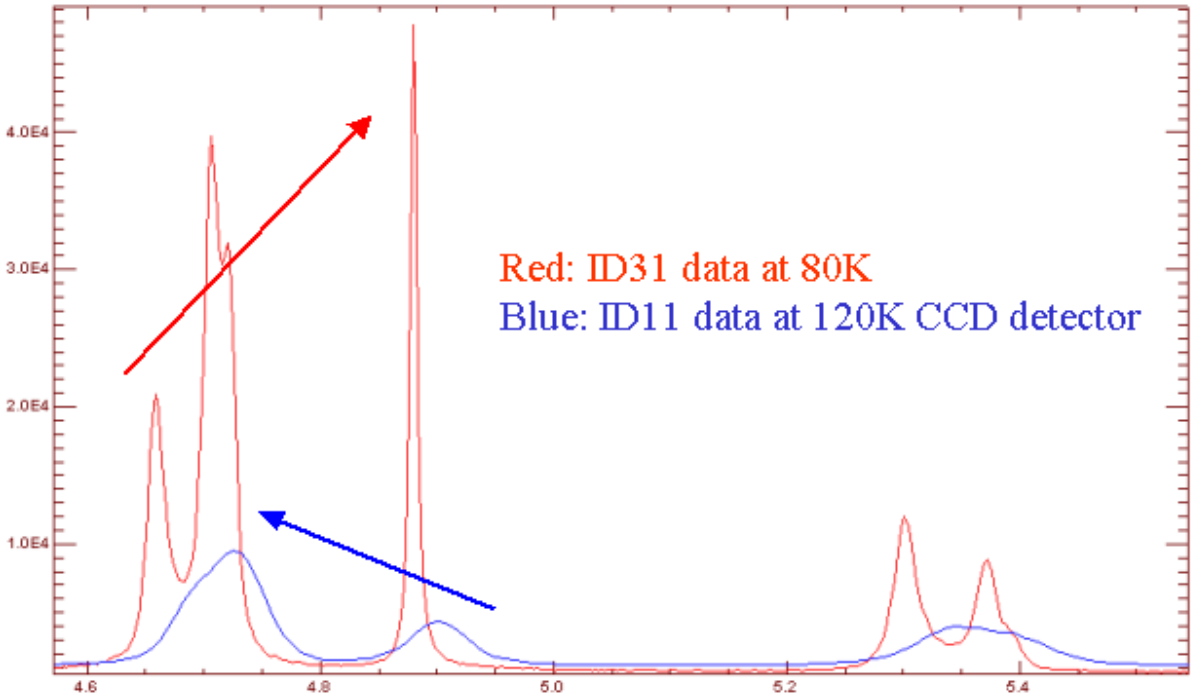
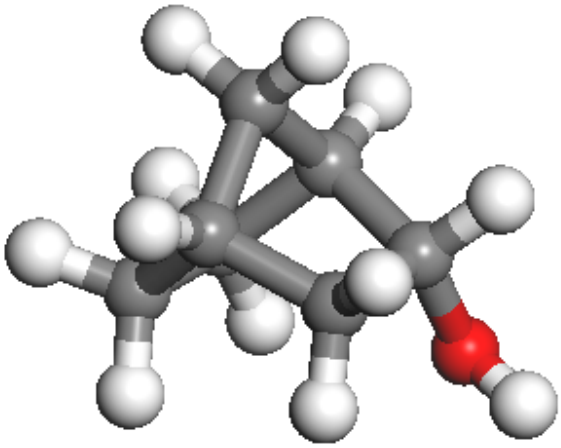


Figure 3-40 Comparison between the data collected with a point detector on ID31 and the area detector on ID11, some peaks have a different relative intensity.

Again it is important to note the big difference in resolution, that can be obtained on the ID31 beamline, that is indispensable in order to find the unit cell.

3.2.6 (±)-endo-norborneol



The solid-state behaviour of the (±)-endo-norborneol, $C_7H_{12}O$, was completely unknown.

At room temperature the sample is solid forming a white sticky powder. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from room temperature phase to the low ordered temperature phases was investigated by recording on beamline ID31 short diffraction patterns of 1 min each. Cooling at the rate of 80 K $hour^{-1}$ we could not observe any transition.

At room temperature the sample was indexed manually as hexagonal with a=10.9331(1) Å c=10.1534(1) Å. The cell was checked also with the SVD procedures and the same cell as already determined manually was found. The space group is $P6_3/m$.

The cell is very similar to the one of the phase I of the (+)-endo-2-norborneol the pure enantiomer but surprising any phase transition occurs.

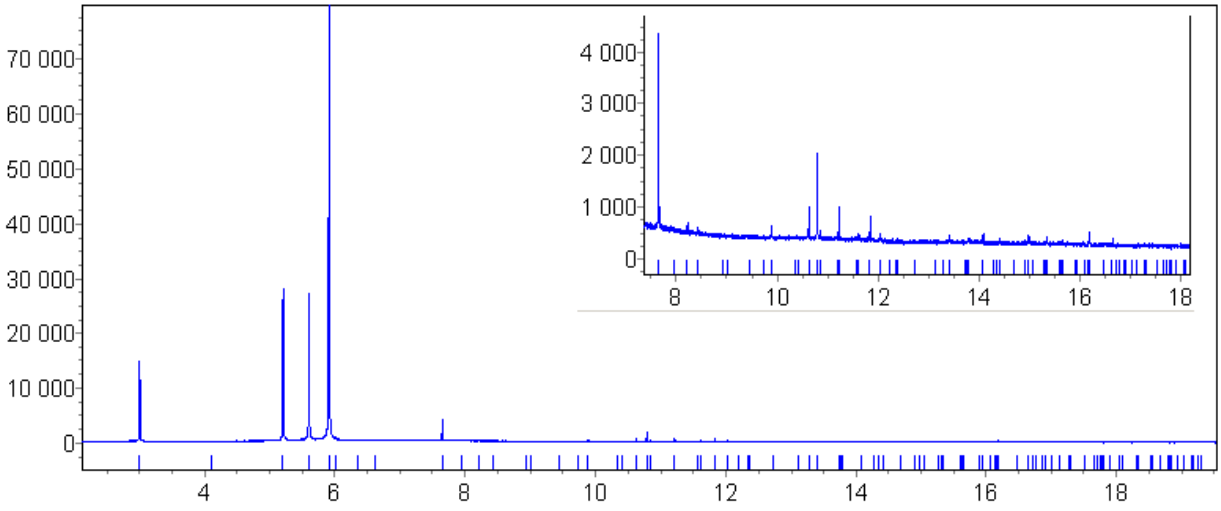
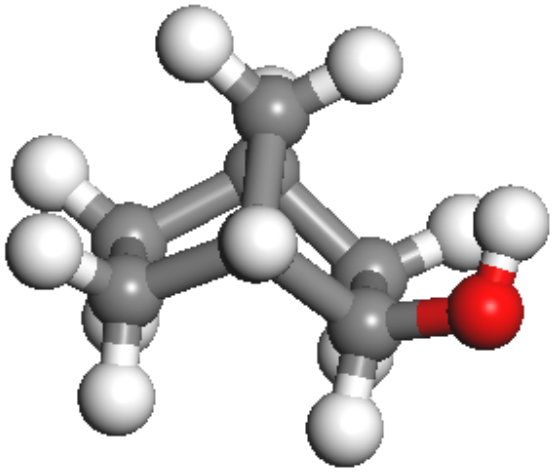


Figure 3-41 Powder diffraction pattern of the disordered hexagonal phase (λ=*0.49617* Å)

3.2.7 (±)-exo-norborneol



The solid-state behaviour of the (±)-exonorborneol, $C_7H_{12}O$, was completely unknown.

At room temperature the sample is solid forming a white sticky powder. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from room temperature phase to the low ordered temperature phases was investigated by recording on beamline ID31 short diffraction

patterns of 1 min each. Cooling at the rate of 80 K hour^{-1} we could observe the transition from the disordered phase I to the ordered phase II occurring at T = 271 K.

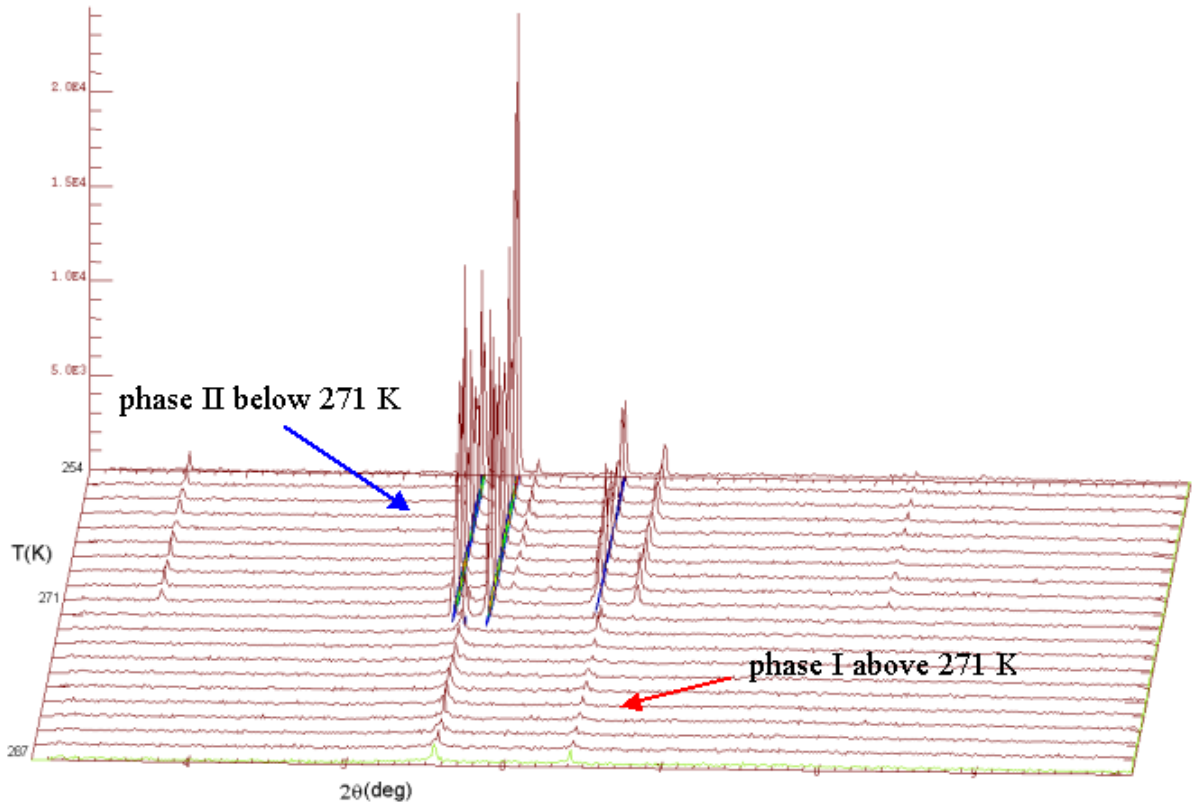


Figure 3-42 Powder diffractograms showing the phase transition

3.2.7.1 Phase I

This phase was already indexed manually as cubic with a=8.9371(2) Å at room temperature. The cell was checked also with the SVD procedures and the same cell as already determined manually was found. The space group is Fm3m.

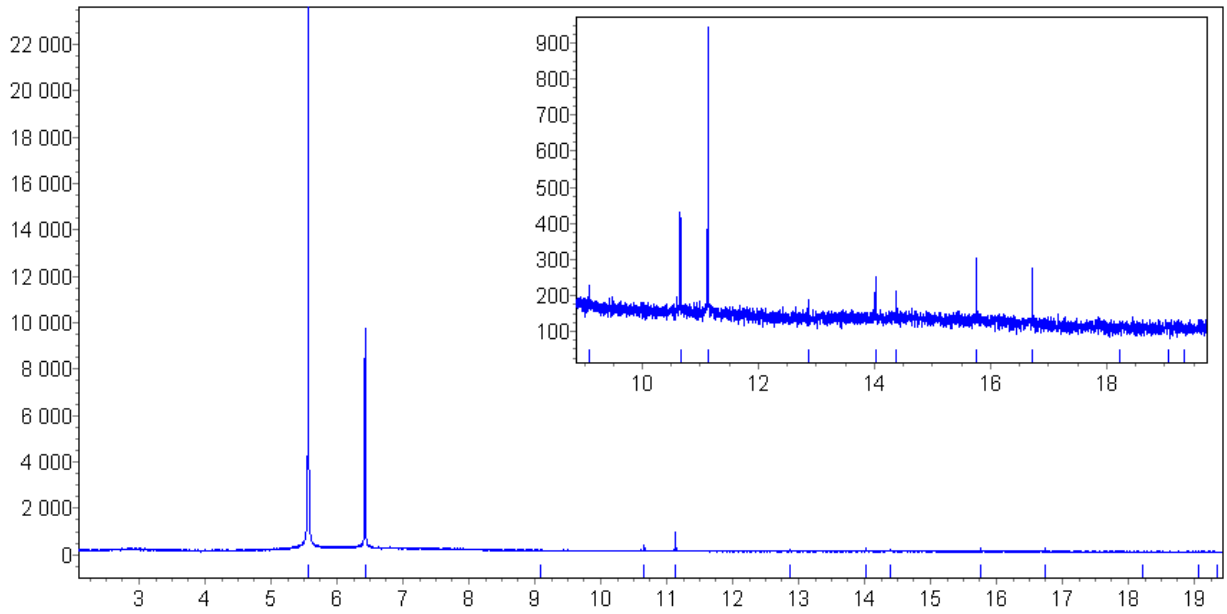


Figure 3-43 Powder diffraction pattern of the disordered cubic phase (λ=*0.50027* Å)

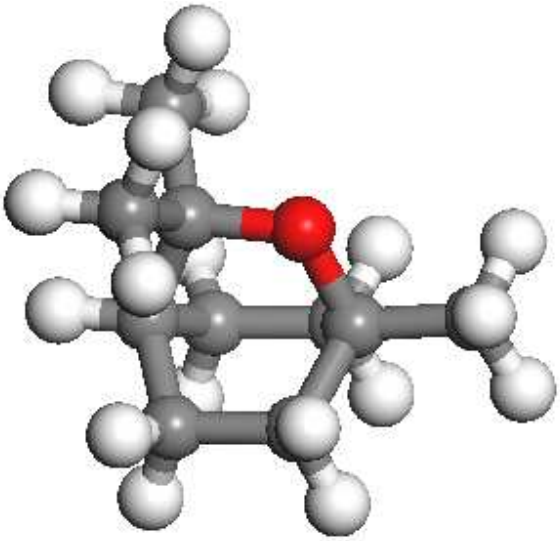
3.2.7.2 Phase II

The diffraction pattern recorded on ID31 at 80 K was indexed from the positions of 20 low-angle peaks using the program FJZONE, with a figure of merit M_{20} of 42, in a orthorhombic cell:

$$a = 12.0866(3)\ \text{Å} \quad b = 8.4233(2)\ \text{Å} \quad c = 19.2488(6)\ \text{Å}$$

The cell volume of 1959.7(1) $Å^3$ indicates 12 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $Pna2_1$, and hence there are 3 independent molecules in the asymmetric unit. In order to solve the structure we used DASH, FOX and TOPAS. None of these attempts was able to find a good solution. As in the case of the (+)-endo-2-norborneol the sample could be affected by strong texture but further measurements must be done to confirm this idea and to see if with the data collected with an area detector is possible to solve the structure.

3.2.8 Eucalyptol



The solid-state behaviour of the Eucalyptol (1-8-Cineole), $C_{10}H_{18}O$, was completely unknown. At room temperature the sample is liquid. The sample was loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of the sample through the phase transitions from room temperature phase to the low ordered temperature phases was investigating by recording on ID31 short diffraction patterns of 1 min each. Cooling at the rate of 80 K hour^{-1} we could observe the transition from the liquid to the ordered phase occurring at T = 250 K. No other transitions where observed. As we will show the low temperature structure wasn't solved owing to strong granularity and preferred orientation too. It looks similar to the case of the (1R)-(+)-nopinone and a more careful investigation of the behaviour of this sample in cooling should be done again using an area detector to see if any plastic phase is present.

The diffraction pattern recorded on the beamline ID31 at 115 K was indexed from the positions of 20 low-angle peaks using the program DICVOL, with a figure of merit M_{20} of 73, in a monoclinic cell:

$$a = 7.80938(6)\ \text{Å}\quad b = 10.59238(5)\ \text{Å}\quad c = 5.97832(5)\ \text{Å}\quad \beta = 110.266(1)^\circ$$

The cell volume of 463.913(1) $Å^3$ indicates 2 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_1$, and hence there is 1 independent molecule in the asymmetric unit.

In order to solve the structure we used DASH, FOX and TOPAS. None of these attempts was able to find a good solution. We measured then the sample on the SNBL and on ID11 but also this data set weren't good enough to solve the structure, due to a strong preferred orientation effect and granularity.

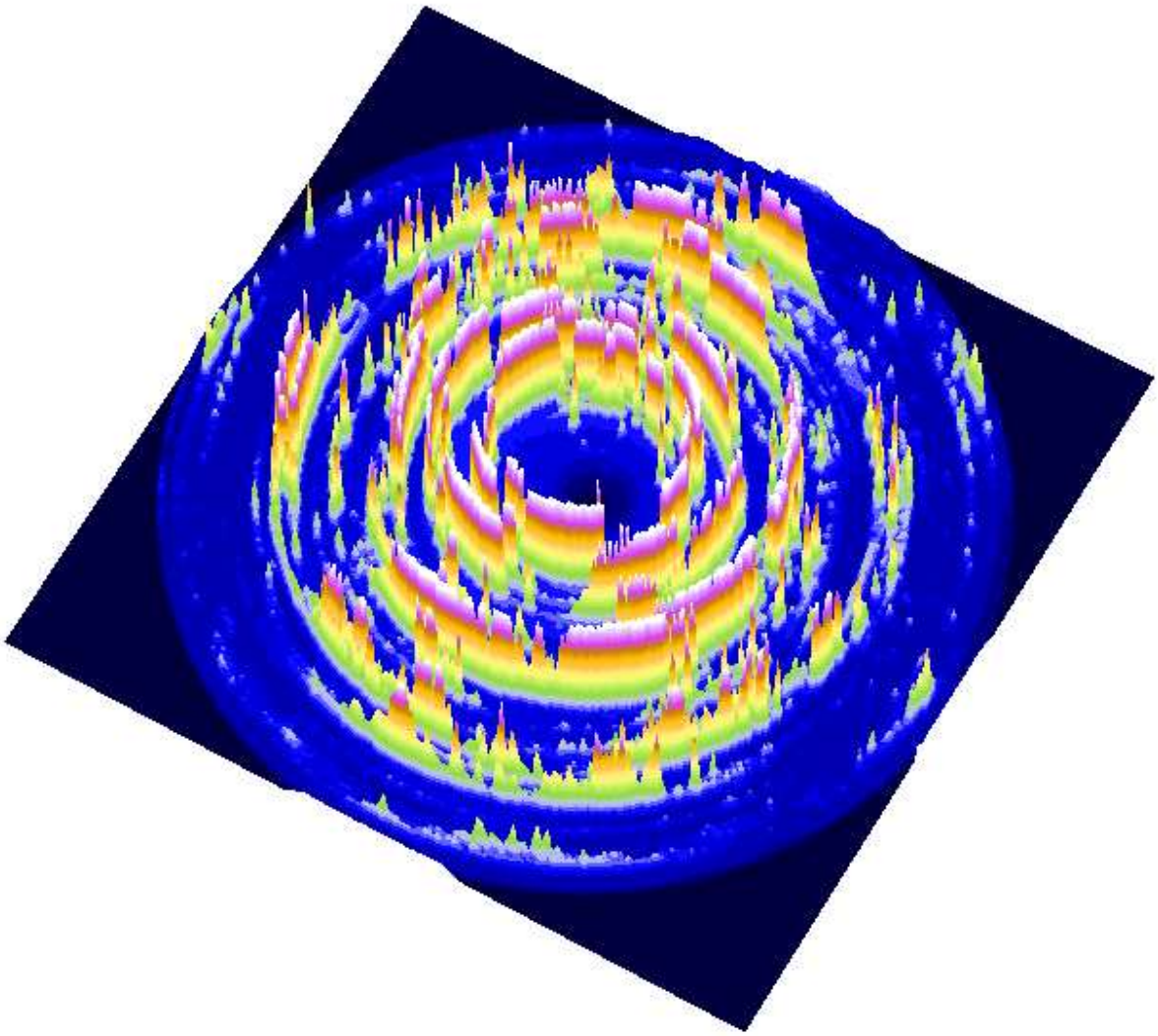
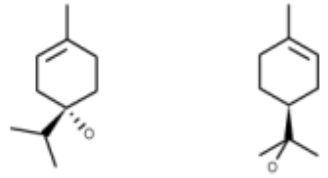


Figure 3-44 3D-plot of the image collected on the SNBL, problem of granularity and preferred orientation

3.2.9 Un-idexed samples



(+)-terpinen-4-ol (-)-a-terpineol

The solid-state behaviour of the (-)-α-terpineol, $C_{10}H_{18}O$, and of the (+)-terpinen-4-ol, $C_{10}H_{18}O$ was completely unknown.

At room temperature both samples are solid forming a white sticky powder. The samples were loaded into a 1-mm-diameter borosilicate thin-walled glass capillary. The evolution of both samples from room temperature to the low temperature was investigated by recording on the ID31 short diffraction patterns of 1 min each. Cooling at the rate of 80 K $hour^{-1}$ we didn't observe any phase transition. For both samples it was not possible to index the unit cell

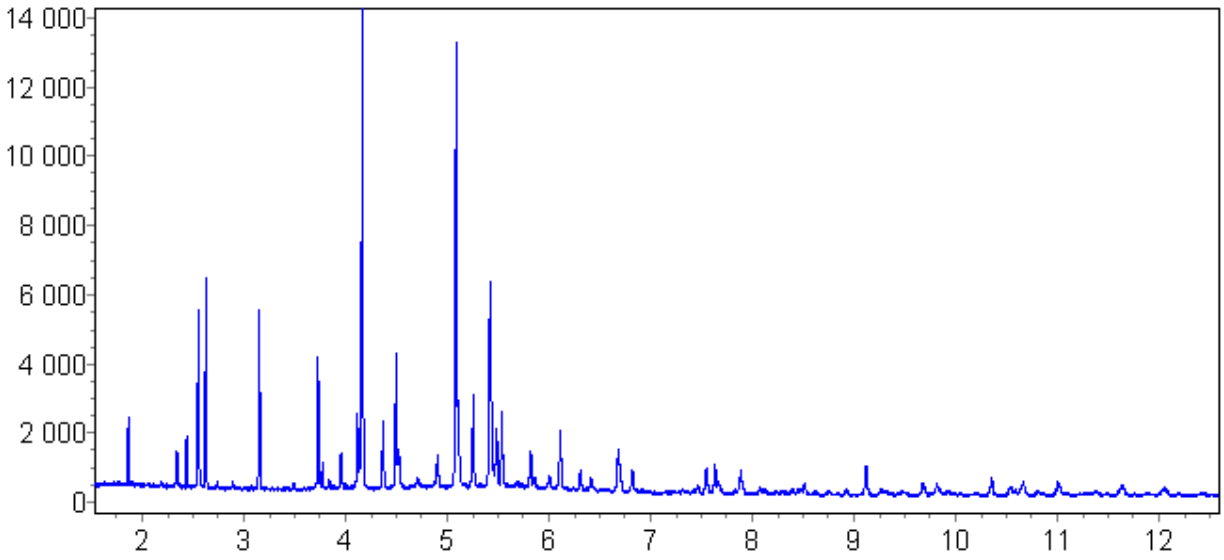
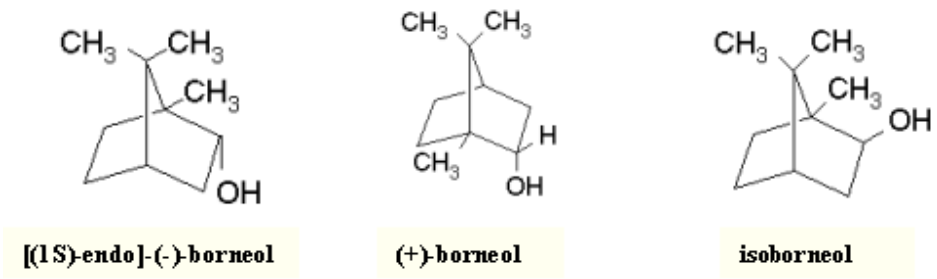


Figure 3-45 Diffractogram of the (-)-α-terpienol, the one of the (+)-terpinen-4-ol looks also similar

Also for this other family of samples, that didn't show any phase transition, it was not possible to find the unit cell:



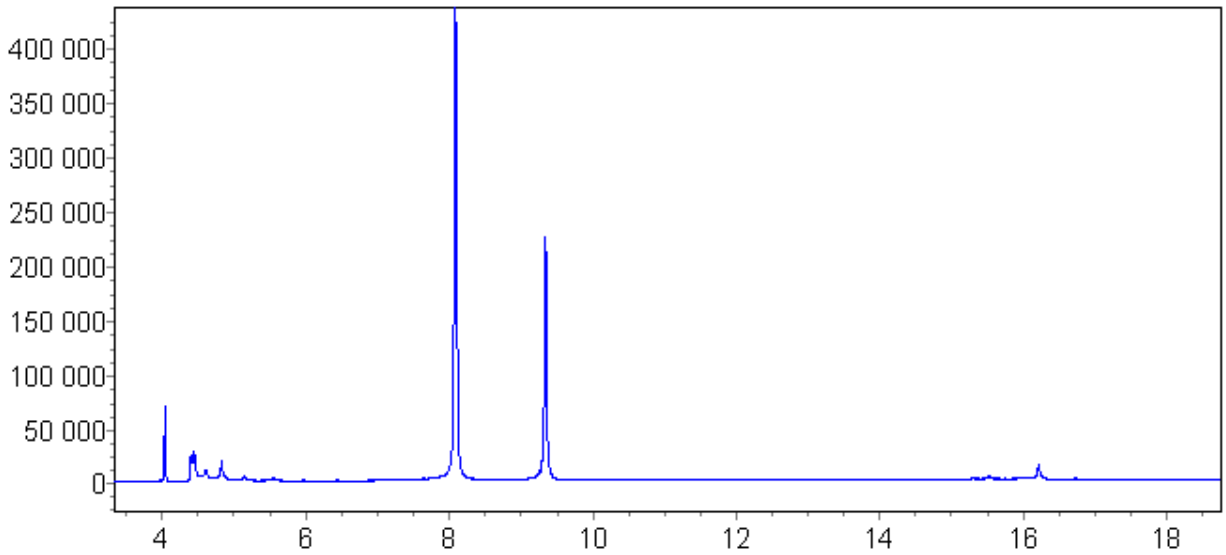
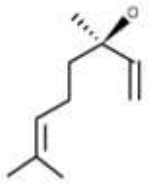


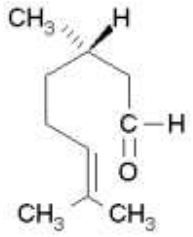
Figure 3-46 Diffractogram of the [(1S)-endo]-(-)-borneol, the (+)-borneol and the isoborneol one look similar

3.2.10 Glassy samples

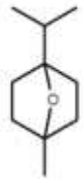
We analyzed the following samples, liquid at room temperature, that refuse to crystallize on lowering the temperature, but stayed in a glassy state. We tried different cooling rate schemes to see if some of these samples would form a crystalline phase but we didn't find any evidence for this. All these samples are viscous oils at room temperature.



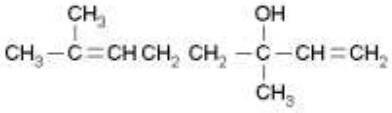
(-)-linalool



(*S*)-(-)-citronellal

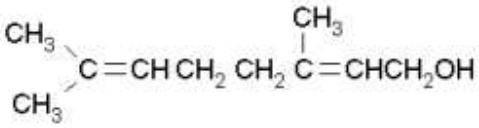


1,4-cineole

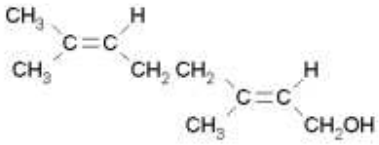


(±)-linalool

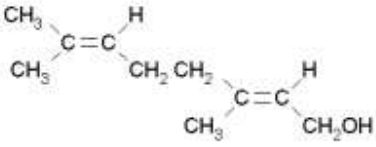
(*S*)-(-)-limonene



nerol

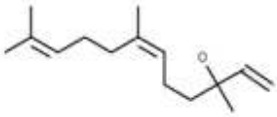


geraniol

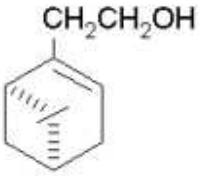


geraniol

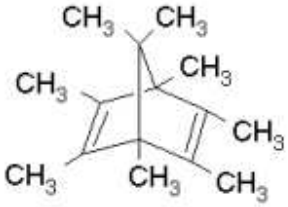
p-cymene



cis-nerolidol



(1R)-(-)-nopol



1,2,3,4,5,6,7,7-octamethyl-bicyclo(2.2.1)hepta-2,5-diene

3.2.11 Summary

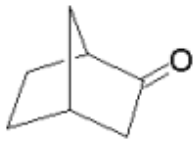
We report here a summary of the result obtained:

7-oxabicyclo[2.2.1]heptane

There are 4 phases:

- Phase I (251-234 K): cubic, Fm3m, a = 8.54690(8) Å at 240K
- Phase II (234-192 K): cubic, Pm3m, a=10.60558 (4) Å at T=194K
- Phase III (191-185K): monoclinic, $P2_1/a$, a=21.852(1) Å, b=10.9699(4) Å, c=9.3331(4) Å, β=90.453(5)° at T=185 K
- Phase IV (<185 K): monoclinic, C2/c, a=13.6378(1) Å, b=10.2596(1) Å, c=24.0470(2) Å, β=102.5217(4)° at T = 50 K

The forces that determine the crystal packing are dipole-dipole type interactions.

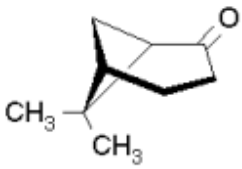


norcamphor

There are 3 phases:

- Phase I (>164 K): hexagonal, $P6_3/mmc$(?), a=6.1785(6) Å, c=9.7062(9) Å at 170K
- Phase II (154-164 K): orthorhombic, Pna21, a=9.1001(3) Å, b=9.3654(3) Å, c=7.3118(2) Å, at T=160 K
- Phase III (<146 K): monoclinic, $P2_1/a$, a=18.464(1) Å, b=12.2889(7) Å, c=10.6852(7) Å, β=90.993(3)° at T= 35 K

The forces that determine the crystal packing are dipole-dipole type interactions, the structure at 35K has not been completely determined due to granularity problem



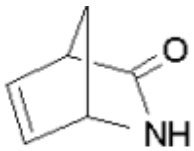
(1R)-(+)-nopinone

There are 2 phases:

- Phase I (260-240K): cubic, unknown sg., a=7.6439 Å, at 255K
- Phase II (<240 K): orthorhombic, $P2_12_12_1$, a=17.65545(6) Å, b=6.96785(3) Å, c=6.52987(2) Å, at T=160 K

The forces that determine the crystal packing are dipole-dipole type interactions

---

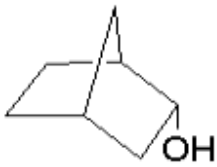


(±)-2-azabicyclo[2.2.1]hept-5-en-3-one

There is 1 phase:

- Phase I: orthorhombic, $P2_12_12_1$, a=6.2179(1) Å, b=14.3605(3Å, c=6.0628(1) Å, at T=80 K

The forces that determine the crystal packing are hydrogen bonding type interactions



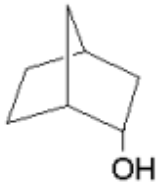
(+)-endo-2-norborneol

There are 2 phases:

- Phase I (>240 K): hexagonal, $P6_3$, a=10.9367(5) Å, c=10.1280(5) Å at room temperature
- Phase II (<240 K): monoclinic, $P2_1$, a=10.7403(2) Å, b=9.7565(1) Å, c=10.6011(2) Å, β=119.301(1)°at T= 35 K

The molecular packing is not determined due to strong texture

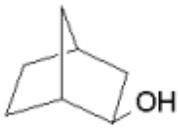
---



(±)-endo-norborneol

There is 1 phase:

- Phase I: hexagonal, $P6_3/m$, a=10.9331(1) Å, c=10.1534(1) Å at 170K

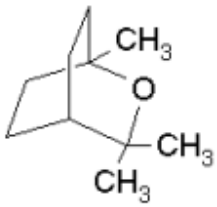


(±)-exonorborneol

There are 2 phases:

- Phase I (>271 K): cubic, unknown sg., a=8.9371(2)Å at room temperature
- Phase II (<271 K): orthorhombic, Pna21, a=12.0866(3) Å, b=8.4233(2) Å, c=19.2488(6) Å, at T=80 K

The molecular packing is not determined due probably to strong texture.



eucalyptol

There is 1 phase:

- Phase I (<250 K): monoclinic, P21, a=7.80938(6) Å, b=10.59238(5) Å, c=5.97832(5) Å, β=110.266(1)° at T= 115 K

The molecular packing is not determined due to strong texture and preferred orientation, maybe the plastic phase exist but more accurate measurement is needed using an area detector

Chapter 4 Neuropeptides

Ce chapitre est centré sur la famille des neuropeptides opiacés. Ces neurotransmetteurs régulent, au sein du système nerveux central, les fonctions sensitives telles que la douleur et la respiration. Le but de notre étude était de déterminer la structure cristalline de la Leu-enkephaline, sous sa forme complète ou fragmentée. Pour cela, nous sommes partis d'une poudre pure lorsque l'échantillon était déjà cristallin ou alors nous l'avons cristallisé en n'utilisant que de l'eau comme solvant, dans le but de se rapprocher de la réalité biologique et de mettre en évidence d'éventuelles différences de structures avec les résultats bibliographiques obtenus à partir de monocristaux. Les structures ont été déterminées grâce à des simulations de Monte-Carlo.

4.1 Neuropeptides

The term peptides (derived from peptones, from Greek `peptos`, *digested)* was introduced in 1902 by Emil Fischer, referring to curious products from the partial degradation of nitrogen-containing nutrients such as egg albumin, milk casein or blood fibrin, known today as proteins. However, the significance and wide role of peptides in all life processes have become apparent only since the 1950s, owing to the continuous development of increasingly sensitive analytical methods. Peptides exert an extremely wide variety of specific functions. They can act as chemical messengers, hormones in the strict sense, intra-or intercellular mediators, highly specific stimulators and inhibitors. Numerous biological activities in the brain and nervous system are carried out by peptides. Peptides make up a large variety of antibiotics and, more recently, peptides that are actively involved in the reactions of the immune system are attracting increasing attention from biomedical researchers and industries.

The field of peptide hormones and neurotransmitters has been active for over three decades. In its infancy, the field was concerned mainly with the isolation and characterization of the peptides. Subsequently, receptor binding and second messenger assays were developed to

estimate the potency of peptide analogues. This resulted in the definition of both receptor agonists and antagonists. With the advances of molecular biology, genes for many of the known peptides were cloned. Agents were defined that alter the peptides gene expression. More recently, the peptides receptors have been cloned and receptor subtypes identified.

Signaling is a neuron chief function. Within the central nervous system (CNS), neurons communicate with one another; in the periphery, nerve cells also signal to glands and muscles. Communication between the various cellular components of the organism is essential for complex animals. In simpler animals and during the early stages in the development of higher forms, communication is mediated by hormones and growth factors that diffuse relatively long distances from secretory cells to target tissues. In the nervous system, chemical messengers are secreted from specialized parts of the neurons, called nerve terminals or nerve endings, to act on receptors in the membrane of neighboring target cells. The targets are close by, not much further than the distance between two adjacent cells. These chemical messengers are called neurotransmitters. Until the beginning of the twentieth century, many physiologists thought that neurons in the CNS were all connected together, forming one great continuous cell or syncytium similar to cardiac muscle. If neurons were structured in this way, there would be no need for chemical transmission between nerve cells, because they would all be connected electrically. Other physiologists thought that neurons were discrete, but made close contacts with other neurons at nerve terminals. Supported by Ramon y Cajal's magisterial histological studies of both the central and peripheral nervous systems, Charles Sherrington named these close contacts synapses (from the Greek, to clasp). The debate was not settled until mid-century when the fine structure of the nervous tissue could be examined. Electron microscopy revealed both the many varieties of Sherringtonian synapses where chemical messengers are released, as well as cytoplasmic bridges between some neurons, called electrical or electrotonic synapses, where neurotransmitters are not required for communication between the cells.

Substances used as neurotransmitters belong to three distinct families: amines, purines and peptides. Often called classical or small molecule transmitters, the members of the amine family include acetylcholine, amino acids and other substances derived from common amino

acids through a few enzymatic steps. All of these transmitters are synthesized in the nerve endings and are then taken up into vesicles for release when the neuron is stimulated. Peptides function as neurotransmitters. Although the number of active neuropeptides is not yet known, there are more than 100 but probably less than 500. Most of these peptides can be grouped into a few distinct families because they share similar amino acid sequences. Like amine transmitters and purines, peptides are also released at nerve terminals from vesicles. Neuroactive peptides are quite short, usually from 3 to 15 amino acids long, compared with an average precursor protein of about 200 amino acids in length. There are several advantages to this apparently wasteful process. One is *amplification*. In some polyproteins, more than one copy of a short amino acid sequence is present. Several identical neuroactive peptides can be cut out from a single precursor molecule. Another advantage is *diversity*. Amino acid sequences for more than one type of neuroactive peptide are present in other polyproteins. Processing of these precursors can result in the production of several different neuropeptides for release from the same vesicle.[60-62]

4.1.1 Opioid peptides

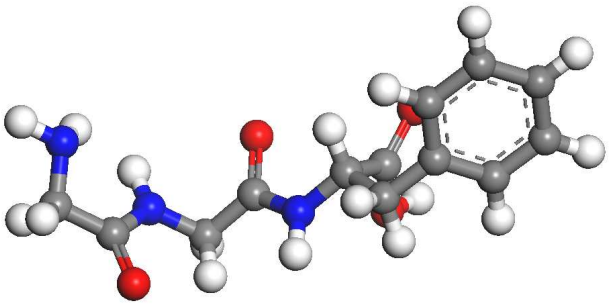
The opiate family of neuropeptides illustrates some of the general principles of peptide transmitters. These peptides all contain the amino acid sequence Tyr-Gly-Gly-Phe. Three types of opioid receptors (δ, κ, μ) have been cloned. They are activated by morphine, the alkaloid from the poppy, to produce relief of pain (narcosis) and many other effects (e.g. sleepiness, euphoria and constipation) that are mediated by the different receptor subtypes distributed on postsynaptic target cells throughout the body. Peptides differ in several important ways from small molecule transmitters. These differences cause neurotransmission by peptides to be slow in onset, to affect targets over a large area, and to last longer than transmission by small molecules. Peptide transmitters are relatively large (even dipeptides or tripeptides are bigger than any small molecule transmitter) and the largest (e.g.insulin) have molecular weights of over 5000 Da. As a result, peptides diffuse slowly. Peptides are not taken back up into neurons or glial cells, nor are they degraded by specific enzymes in the synaptic cleft. Diffusion is the chief mechanism for removing peptide transmitters.

In 1975, Huges and Kosterliz[63] discovered the opioid penta-peptides Met-enkephalin and Leu-enkephalin, which started exstensive research on their biology and chemistry. The ratio of the two enkephalins depended on the species in which they were found. These neurotransmitter regulate sensory functions including pain and control of respiration in the central nervous system by binding to the G-protein coupled μ- and δ-opioids receptors, respectively. The non-specific receptor affinity of the enkephalins has been associated with their large conformational freedom. Therefore, it has been a great challenge to determine the bioactive conformation of the enkephalins. The structural determination of the membrane bound opioid receptor is extremely difficult and short linear peptides agonists are very flexible. Leu-enkephalin has been found to exist, from crystallographic studies, predominantly in three different conformations; extended, single-bend and double-bend depending on which bio-mimetic media was used to grow some nice big single crystals. Two different extended conformations where found by crystallizing the sample using N,N-dymethylformamide (DMFA)[64] or ethanol solutions[65]. The single-bend conformation was obtained crystallizing the sample from an aqueous methanol solution[66]. The double-bend conformation was obtained by slow cooling of an aqueous solution[67]. During the last few years attention has been paid to the fact that the agonist-receptor interaction takes place inside a hydrophobic cavity, and that the agonist, in moving from the transport medium to the receptor, crosses different phases characterized by increasing ordered water (at the membrane/water interface), decreasing dielectric constant and reduced hydrogen-bonding ability. A few studies have been also done on Leu-enkephalin fragments, Try-Gly-Gly-Phe, Gly-Gly-Phe-Leu[68] and Tyr-Gly-Gly[69, 70].

Starting from those result the aim of our study was to try to determine the crystal structure of the Leu-enkephalin and some of the fragments, just starting form the pure powder in the case that it was already crystalline or just crystallizing the sample simply using water without the use of any solvent in order to be maybe closer to biological reality and see if any structural differences were present with respect to the results found in literature. The samples analyzed were the following:

- Gly-Gly-Phe
- Gly-Phe-Leu
- Tyr-Gly-Gly-Phe
- Leu-enkephalin

4.2 Glycine-Glycine-Phenylalanine



The crystal structure of H-Gly-Gly-Phe-OH, $C_{13}H_{17}N_3O_4$ was unknown. The as-received sample (Bachem) was measured on the beamline ID31 at room temperature. The sample is crystalline but it was not possible to index the diffractogram. The sample seems to be disordered and not well crystalline looking at the broadening of the peaks. That is the reason why the indexing fails.

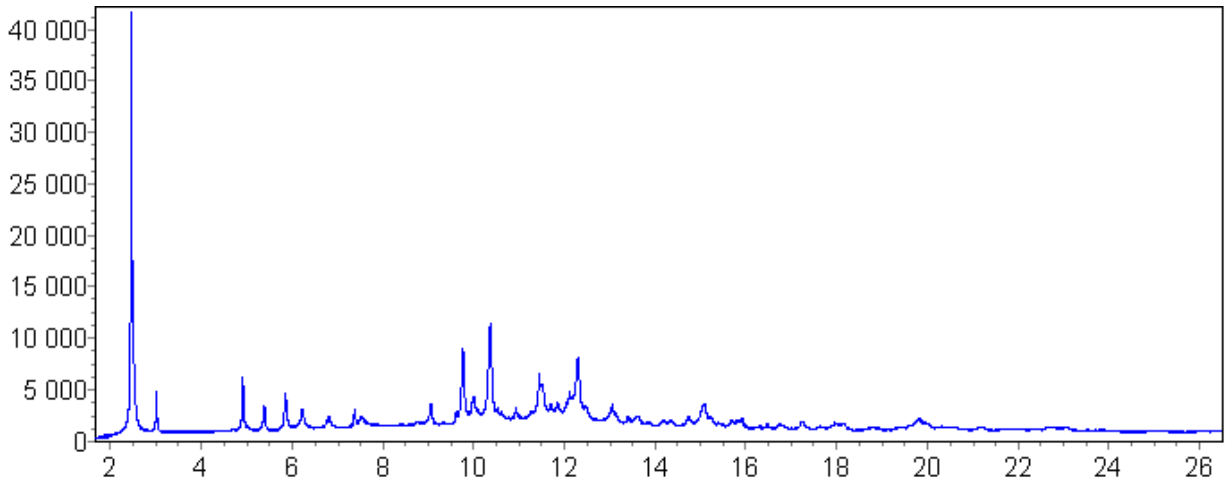


Figure 4-1 Diffractogram recorder on ID31 at room temperature of the sample H-Gly-Gly-Phe-OH as-received

Then we re-crystallized the sample using pure water and we obtained a nice powder that we measured again at room temperature.

The diffraction pattern was indexed from the positions of 20 low-angle peaks using the program DICVOL, with a figure of merit M_{20} of 50, in an orthorhombic cell:

$$a = 30.3902\ (2)\ \text{Å}\quad b = 10.25971(8)\ \text{Å}\quad c = 4.83972\ (4)\ \text{Å}$$

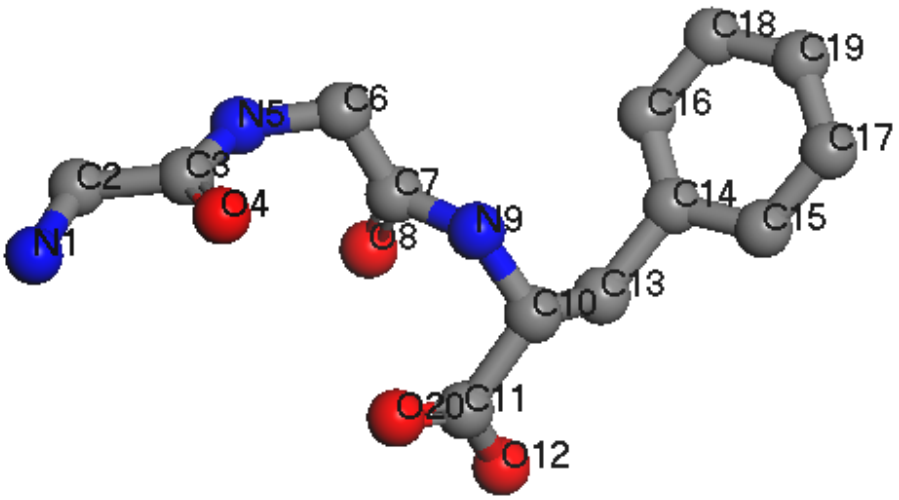
This cell doesn't fit the diffractogram of the as-received sample, thus the unit cell for the as-received compound is different.

The cell volume of 1509.00(1) $Å^3$ indicates 4 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_12_12_1$, and hence there is 1 independent molecule in the asymmetric unit. The structure was solved with the program FOX. As the sample has been crystallized from water there is also an unknown number of water molecules to be localized, we performed different simulations with different amounts of water. With simulation it was difficult to understand if there were 2 or 3 independents water molecules in the asymmetric unit; only in the refinement stage it was possible to understand that there were actually 2. The structure was refined using the program TOPAS. We also performed simulations with TOPAS and found the same solution as with FOX. In the first stage of the refinement only the orientational and positional parameters were left free to vary. At the end the rigid body was almost completely relaxed, everything was free to

vary except the hydrogen atoms, which were kept fixed in the right geometry. Thanks to the good quality of the data we didn't use any restraints. As shown in Table 4-1, only 4 bond length values strongly disagree with the values found in the work of Engh et al.[71], an article written for protein structure refinement where bond-length and bond-angle parameters are derived from a statistical survey of X-ray structure of small compounds from the Cambridge Structural Database. In the same table we also show the values reached imposing the values found in the article, as restrained distances, with a tolerance value of ± 0.01 Å and weight 1.
This structure was refined as if it was a single crystal structure. No planar restraints where imposed on the peptide bonds and on the phenyl ring of the phenylalanine and the final torsion angles coming from the refinement are sensible, only the phenyl ring is a little bit distorted

C6	N5	C3	C2	181.1(7)°	tors. pept. bond
C10	N9	C7	C6	180.4(9) °	tors. pept. bond
C19	C17	C15	C14	351(1) °	tors. hexag. phe

The peaks were described by a PVII peak shape function. The temperatures factors were refined isotropicaly and individually for each atom except the hydrogen atoms that were constrained to the same value. The final refinement, without the use of any additional restraints, converged to R_{wp} = 9.2 %, R_p= 7.7 % and χ^2=1.7



Vector	Length (Å) article Engh	Length (Å) no restr.	Length (Å) with restr.
N1_C2	1.489(33)	1.527(8)	1.489(8)
C2_C3	1.516(18)	1.529(10)	1.516(1)
C3_O4	1.231(20)	1.270(9)	1.239(1)
C3_N5	1.329(14)	1.324(9)	1.329(1)
N5_C6	1.451(16)	1.446(8)	1.449(8)
C6_C7	1.516(18)	**1.405(9)**	1.514(1)
C7_O8	1.231(20)	**1.338(8)**	1.248(9)
C7_N9	1.329(14)	1.350(8)	1.328(8)
N9_C10	1.458(19)	1.430(7)	1.456(8)
C10_C13	1.530(20)	1.512(8)	1.528(9)
C13_C14	1.502(23)	**1.572(9)**	1.503(9)
C14_C15	1.384(21)	1.409(9)	1.396(1)
C15_C17	1.382(30)	1.417(1)	1.389(1)
C17_C19	1.382(30)	1.410 (9)	1.391(1)
C19_C18	1.382(30)	1.397(7)	1.392(8)
C18_C16	1.382(30)	1.403(1)	1.399(1)
C16_C14	1.384(21)	1.400(1)	1.395(1)
C10_C11	1.525(21)	**1.613(1)**	1.528(1)
C11_O20	1.231(20)	1.237(9)	1.233(9)
C11_O12	1.304(22)	1.268(8)	1.300(9)

Atom	B_{iso} no rest. (Å^2)	B_{iso} restr. (Å^2)
N1	1.3(2)	1.7(2)
C2	1.7(2)	1.2(2)
C3	3.0(3)	3.7(3)
O4	3.2(1)	3.4(2)
N5	2.8(2)	2.6(2)
C6	2.4(2)	1.9(2)
C7	2.1(2)	2.4(2)
O8	2.8(1)	2.5(1)
N9	1.4(1)	1.1(2)
C10	1.6(2)	2.4(2)
C11	1.3(2)	2.4(3)
O12	2.7(1)	2.0(1)
C13	2.7(2)	4.0(3)
C14	1.5(2)	1.5(2)
C15	2.2(2)	1.9(2)
C16	3.9(3)	3.9(3)
C17	4.2(3)	3.7(3)
C18	5.0(3)	4.9(3)
C19	3.7(3)	4.5(3)
O20	2.4(1)	1.9(1)
O21	5.7(2)	4.9(2)
O22	3.5(1)	3.7(1)
H	2.5(6)	1.5(6)

Table 4-1 Refined bond lengths values and B_{iso} values with and without restrain

The final refinement using restraints converged to R_{wp} = 9.4%, R_p= 7.9% and χ^2=1.7.

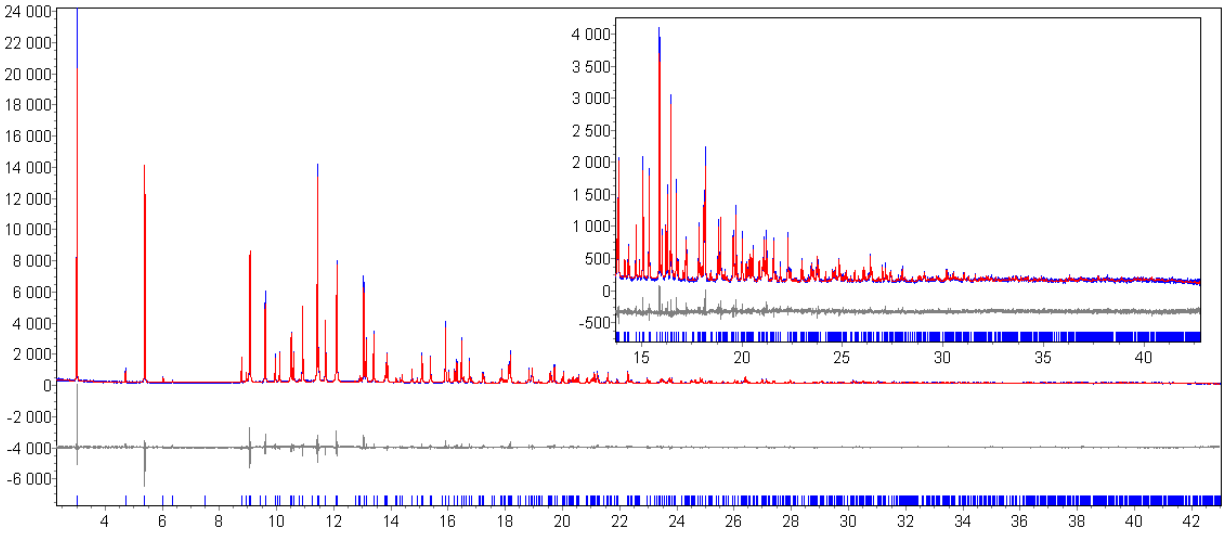


Figure 4-2 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement of the H-Gly-Gly-Phe-OH, crystallized form water, using restraints.

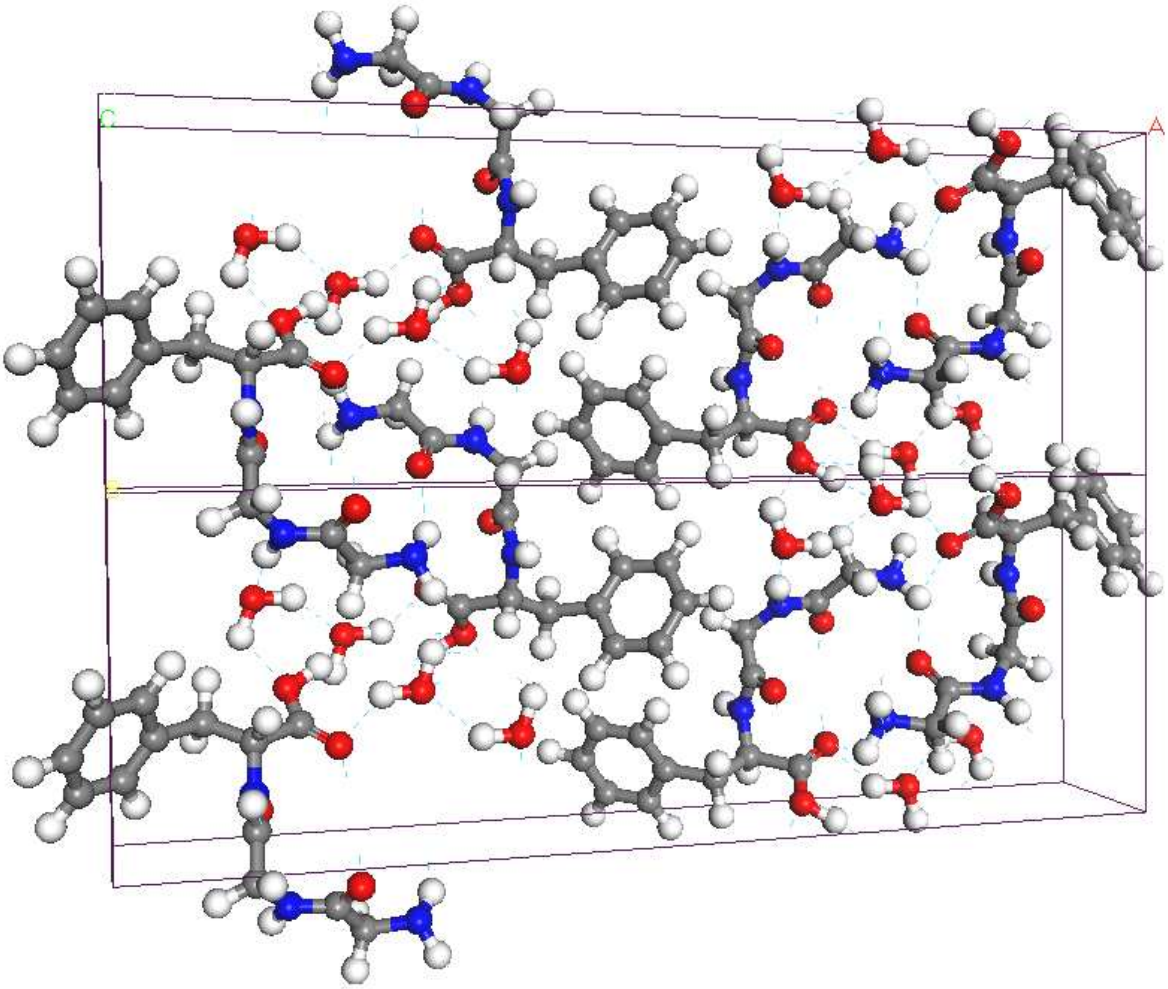


Figure 4-3 Packing of the molecules and interaction with the waters molecules

The forces that determine the crystal packing are hydrogen bond type. The interaction are intermolecular between the water molecules that create a bridge between the two gly-gly-phe molecules and also directly between the two molecules from the atoms N1-O4 and N9-O8.

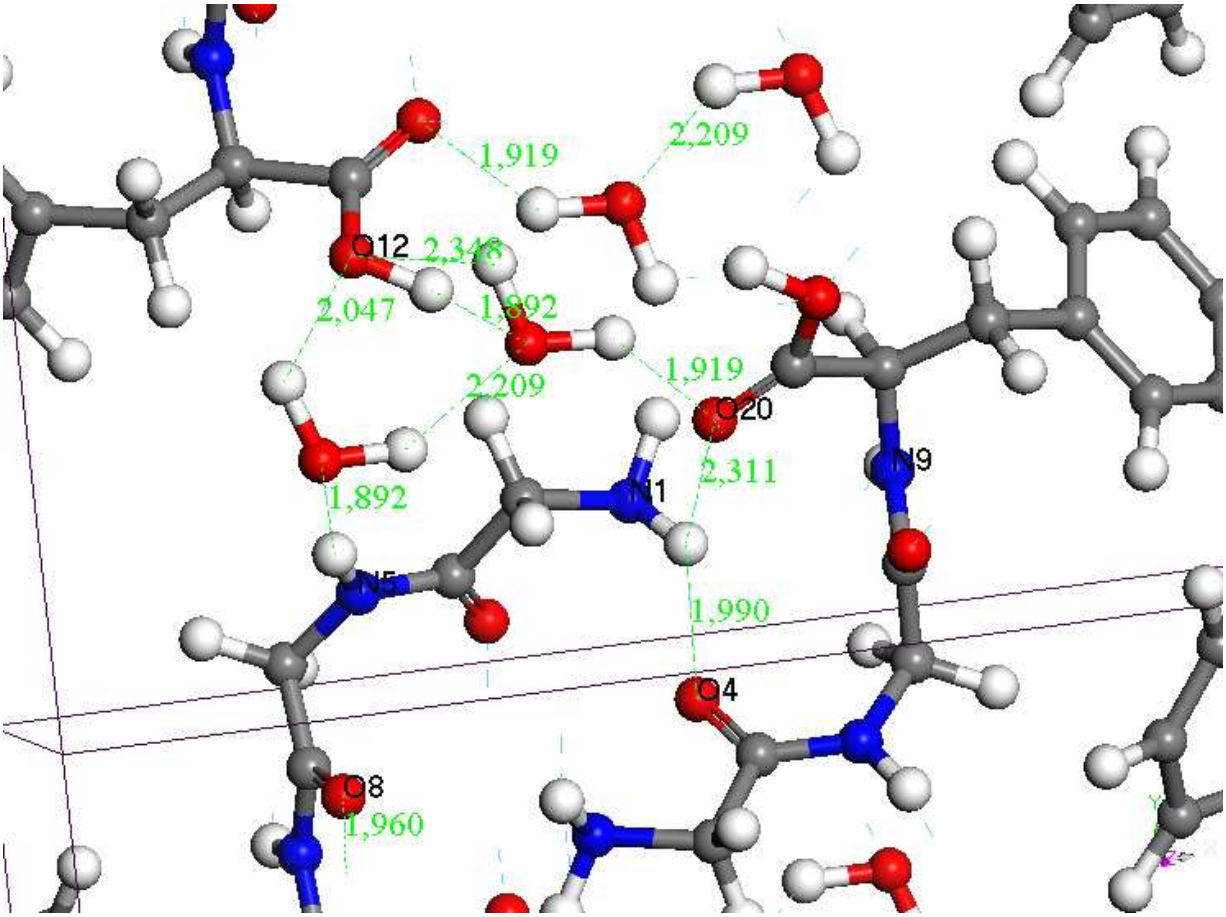


Figure 4-4 Hydrogen bonds

N1	x	0.27265	y	0.19246	z	0.34474
C2	x	0.22604	y	0.21546	z	0.42638
C3	x	0.19771	y	0.12501	z	0.25417
O4	x	0.21391	y	0.05546	z	0.07132
N5	x	0.15494	y	0.12382	z	0.31179
C6	x	0.12405	y	0.0405	z	0.16717
C7	x	0.12601	y	-0.09653	z	0.28267
O8	x	0.13152	y	-0.11533	z	0.53528
N9	x	0.12272	y	-0.20421	z	0.1316
C10	x	0.12401	y	-0.33966	z	0.22151
C11	x	0.16576	y	-0.35967	z	0.3921
O12	x	0.15934	y	-0.44692	z	0.58281
C13	x	0.07918	y	-0.38197	z	0.33299
C14	x	0.03947	y	-0.35943	z	0.15381
C15	x	0.02731	y	-0.46957	z	0.00248
C16	x	0.01066	y	-0.25628	z	0.20418
C17	x	-0.0083	y	-0.45761	z	-0.17537
C18	x	-0.0267	y	-0.2431	z	0.03794
C19	x	-0.0333	y	-0.34478	z	-0.14806
O20	x	0.20033	y	-0.3027	z	0.33499
O21	x	0.21929	y	-0.56887	z	-0.05389
O22	x	0.13044	y	-0.68652	z	-0.24895
H23	x	0.14197	y	0.18352	z	0.48485
H24	x	0.13519	y	0.02389	z	-0.05498
H25	x	0.08996	y	0.08681	z	0.17181
H26	x	0.11902	y	-0.19721	z	-0.09609
H27	x	0.13393	y	-0.40406	z	0.03995
H28	x	0.16577	y	-0.28256	z	0.54546
H29	x	0.19112	y	-0.49408	z	0.63564
H30	x	0.07096	y	-0.32126	z	0.5234
H31	x	0.08054	y	-0.48923	z	0.39385
H32	x	0.04725	y	-0.56326	z	0.02128
H33	x	0.01836	y	-0.18262	z	0.37378
H34	x	-0.0179	y	-0.54233	z	-0.31549
H35	x	-0.0513	y	-0.16158	z	0.08026
H36	x	-0.0634	y	-0.32923	z	-0.28451
H37	x	0.20975	y	-0.46444	z	-0.055
H38	x	0.2556	y	-0.57708	z	-0.03608
H39	x	0.11884	y	-0.58393	z	-0.24808
H40	x	0.16695	y	-0.68832	z	-0.25242
H41	x	0.28964	y	0.13683	z	0.50997
H42	x	0.28945	y	0.28747	z	0.31459
H43	x	0.21595	y	0.32013	z	0.37909
H44	x	0.22184	y	0.19348	z	0.65576

Table 4-2 Atomic fractional coordinates derived form the rigid body

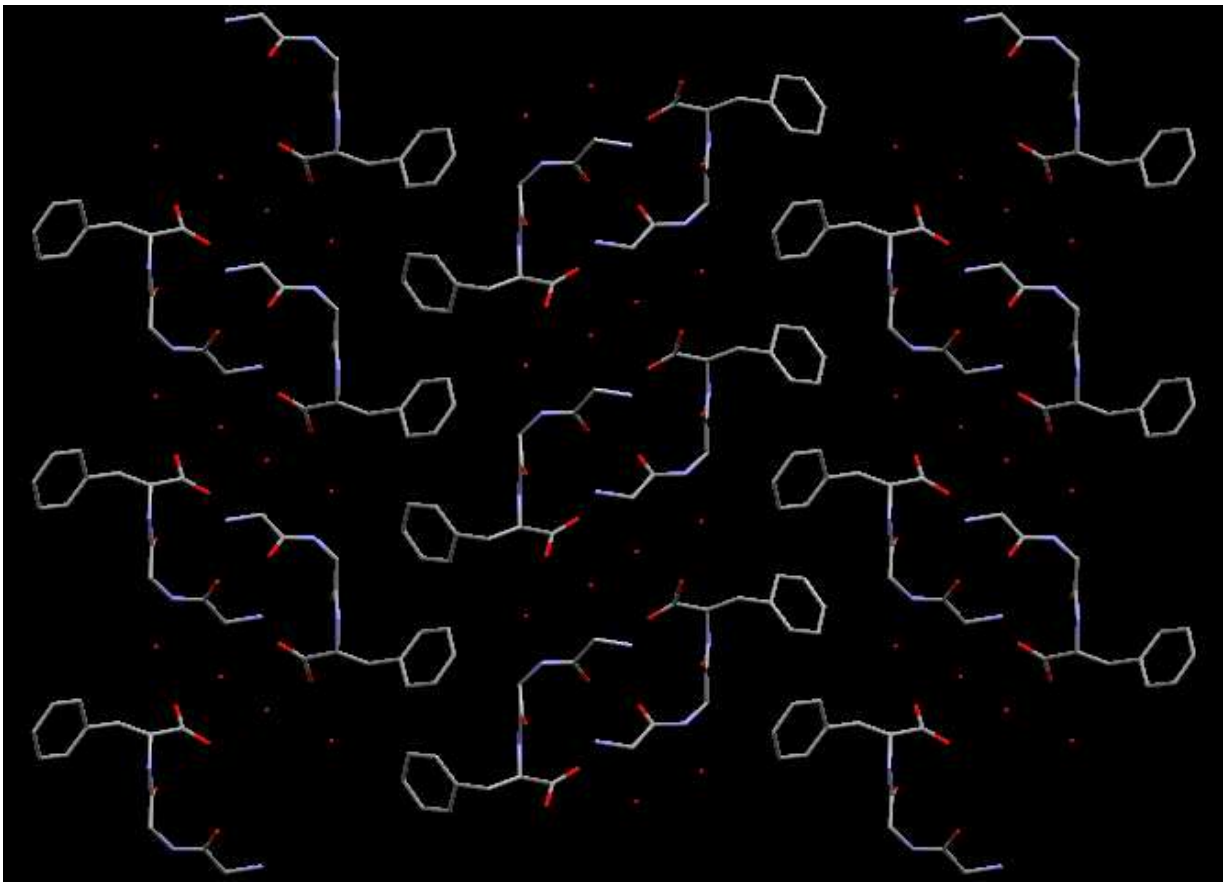
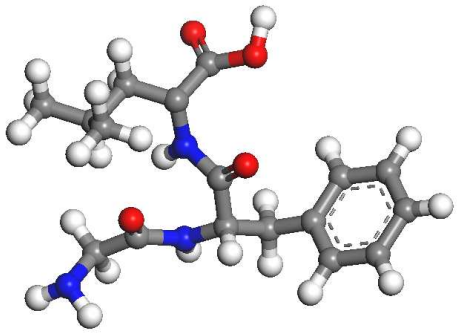


Figure 4-5 Another view of the packing of the molecules

4.3 Glycine-Phenylalanine-Leucine



The crystal structure of H-Gly-Phe-Leu-OH, $C_{17}H_{24}N_3O_4$ was unknown. The as-received sample (Bachem) was measured on beamline ID31 at room temperature. The sample is crystalline. The diffraction pattern indexed from the positions of 20 low-angle peaks using the program DICVOL, with a figure of merit M_{20} of 50, in a monoclinic cell:

$$a = 20.0024(8)\text{Å} \quad b = 4.8738(1)\ \text{Å} \quad c = 10.2778(4)\ \text{Å}\ \beta = 103.940(3)^\circ$$

The cell volume of 972.51(4)Å^3 indicates 2 molecules per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_1$, and hence there is 1 independent molecule in the asymmetric unit. The structure has been solved and refined with TOPAS. The sample is disordered but an approximate solution has been found. Comparing this data with the data collected on the Gly-Gly-Phe it is possible to see how peaks are broader, an indication of the disorder.

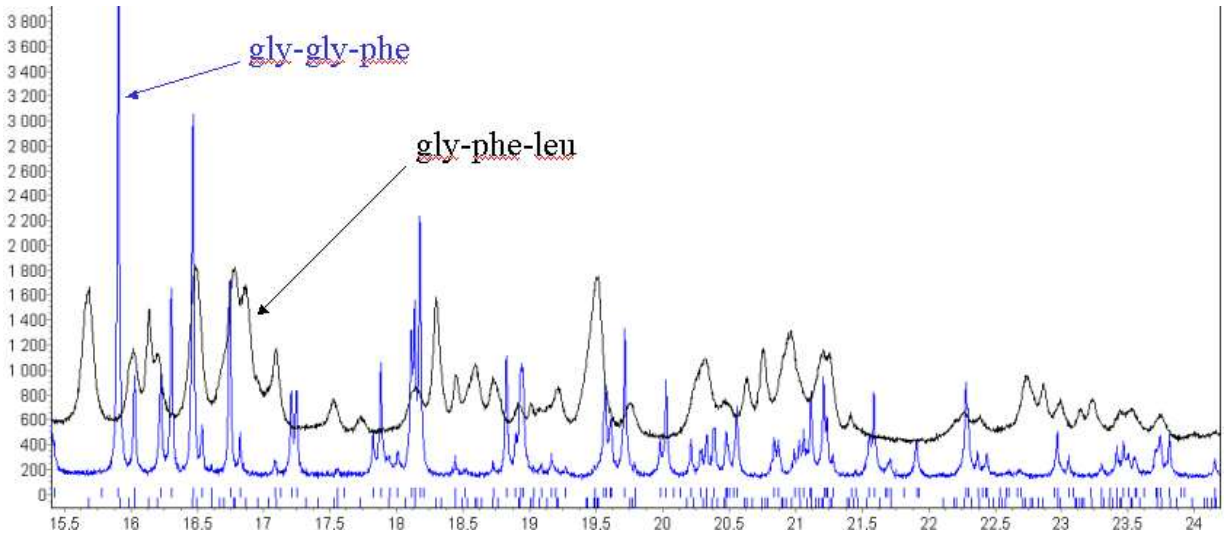


Figure 4-6 Comparison between the gly-gly-phe sample crystallized form water and the gly-phe-leu as-received

Because of the disorder we didn't use strong restraints but we set limits in the bond and angle values to avoid ending up with a non-meaningful structure.

The final refinement converged to R_{wp} = 11.3 %, R_p= 8.4 % χ^2=7.1. An overall temperature factors was refined for the whole molecule reaching the final value of B=3.0(1) Å^2. The position of the center of mass of the molecule and the conformation assumed in the crystalline state can be roughly determined but we can't have any information about bond lengths and angle values. The data are fitted well for reflections bigger than 3 Å in d-spacing; reflections of lower d-space are instead not fitted perfectly.

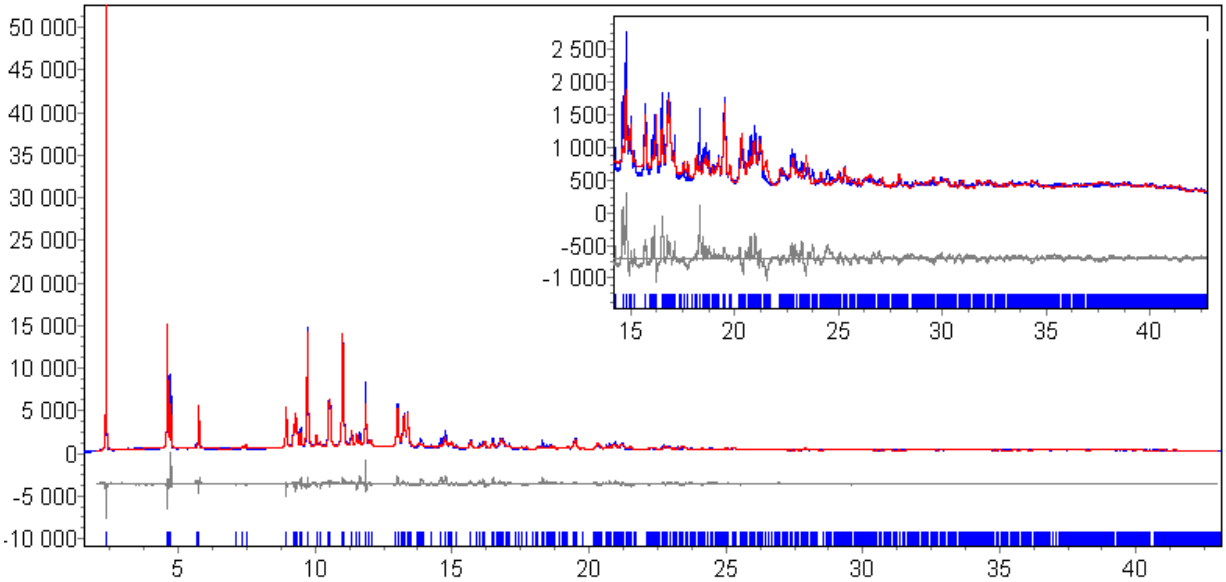


Figure 4-7 Observed (+,blue), calculated(-, red) and difference plot of the final Rietveld refinement of the H-Gly-Phe-Leu-OH

To have a representation of what kind of information we can obtain from this refinement, we computed a MEM map, Figure 4-8, with the program PRIMA starting from a uniform prior, just using the F_{obs} corresponding to d-spacings bigger than 3 Å. We can do this computation because in the MEM calculation there is not the truncation problem that occurs with the Fourier synthesis. MEM is very useful in reconstructing images from incomplete data.

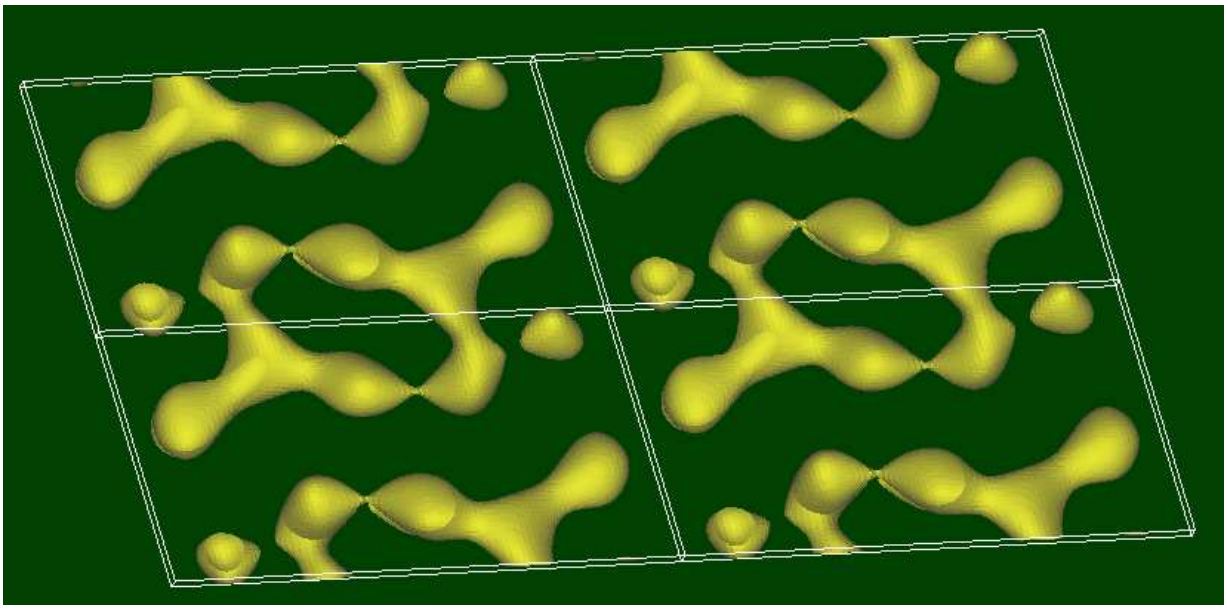


Figure 4-8 MEM map computed with the program PRIMA, starting from an uniform prior, just using the F_{obs} corresponding to d-spacing bigger than 3 Å.

In Figure 4-9 is represented the same result with the atomic model superimposed.

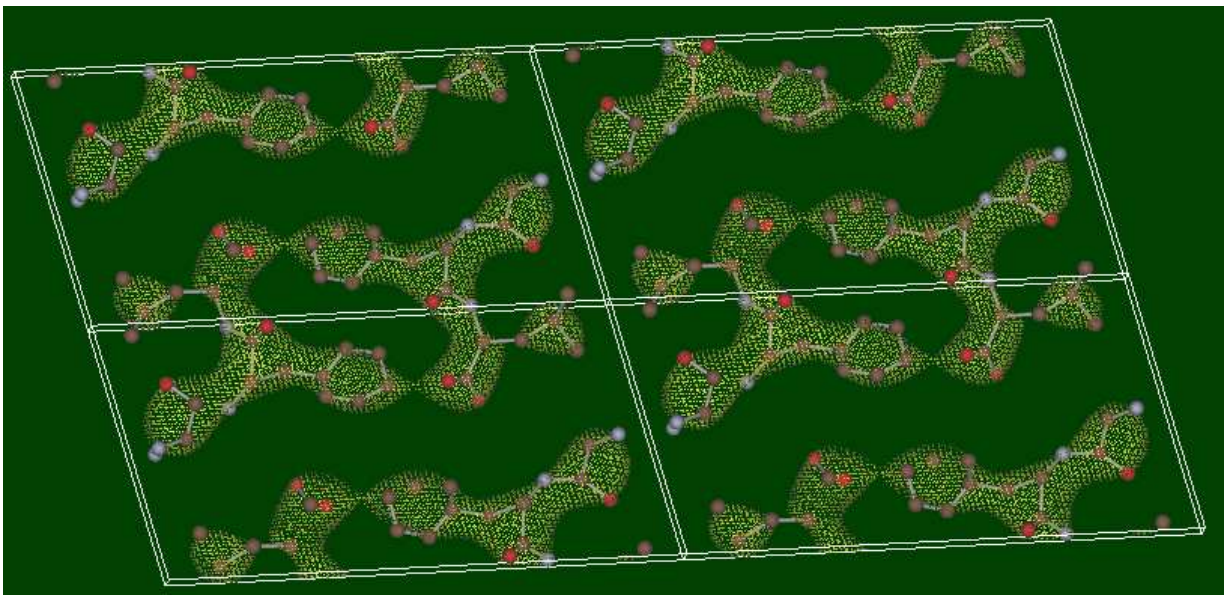


Figure 4-9 MEM map computed with the program PRIMA, starting from an uniform prior, just using the F_{obs} corresponding to d-spacing bigger than 3 Å, with the atomic model superimposed.

We then performed another MEM using this time all the F_{obs}

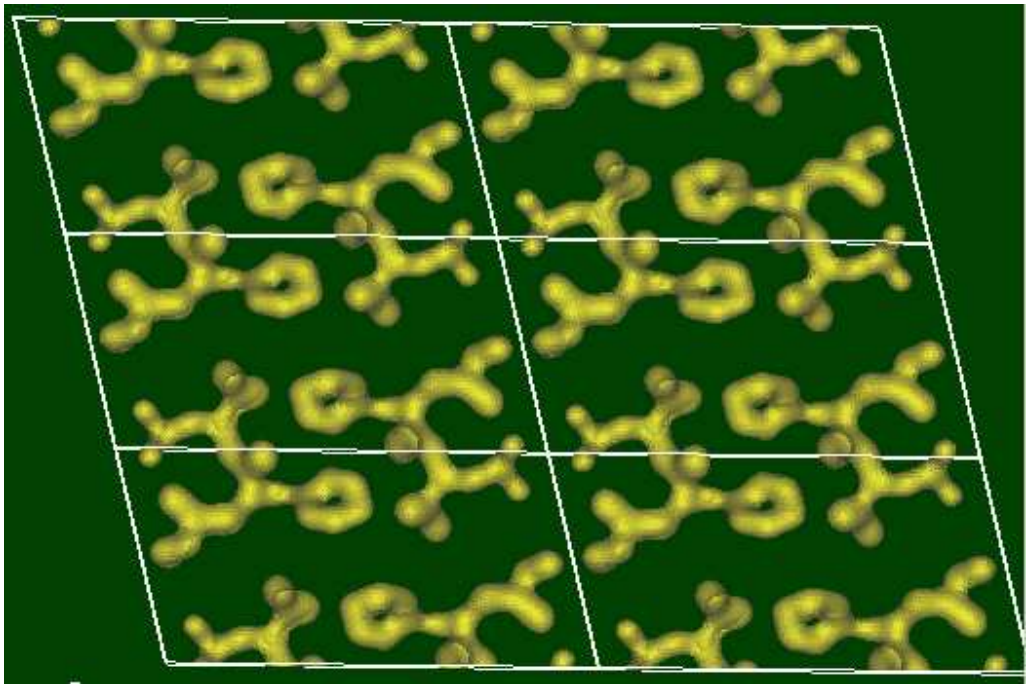
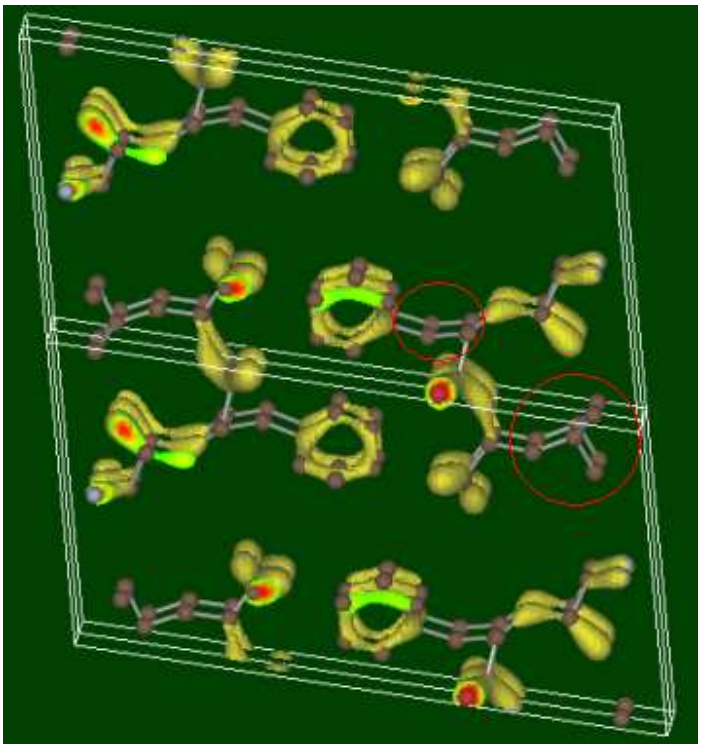


Figure 4-10 MEM map computed with the program PRIMA, starting from an uniform prior, just using all F_{obs}

With these maps maps although providing an alternative in representing the structure with an atomic model, the improvement with respect to the initial model is minimal. To extract the F_{obs} we need the atomic coordinates and then also the MEM maps are model biased. This is the best result we could obtain from the data collected but it is to be taken into consideration that the model is almost right but some torsion angles could be slightly different from the representation we proposed. Looking at the disposition of the phenyl ring we can see some similarity with the Gly-Gly-Phe structure (Figure 4-3).

We tried to use this map in order to see the regions where the density was lower, a possible indication of the disorder.



The side chains of the leucine and the phenylalanine have the lowest amount of density. We then performed different simulations in TOPAS in order to try to describe the disorder in those regions but without success, the fit improved but not in a significant way.

We also tried to re-crystallize the sample from water but the sample is not soluble in pure water and after water evaporation, the diffractogram is the same as for the pure sample.

4.4 Tyrosine-Glycine-Glycine-Phenylalanine

The pure sample was amorphous and the sample is soluble in water but doesn't crystallize. We tried different solution concentrations and different evaporation rates but with no success.

4.5 Leu-enkephalin

Leu-enkephalin, H-Tyr-Gly-Gly-Phe-Leu-OH, $C_{28}H_{37}N_5O_7$ has been found to exist predominantly in three different conformations; extended, single-bend and double-bend depending on the crystallization conditions.

The pure as-received sample is amorphous.

We then crystallized it starting from pure water. The first attempts where done by letting a solution formed by 10 mg of sample and 1 ml of water evaporate in a desiccator. We obtained a powder with also some micro crystals. We measured then the powder at room temperature on ID31 and the micro crystals on ID11.

The cell found on ID11 matched the one form the powder diffractogram obtained with the SVD procedures present in TOPAS. The cell is monoclinic:

$$a = 8.3677(2)\text{Å} \quad b = 11.2142(3)\ \text{Å} \quad c = 31.899(1)\ \text{Å} \quad \beta = 97.675(3)^\circ$$

The cell volume of 2966.57 Å^3 indicates 4 molecules in per unit cell. The systematic absences in the diffraction patterns indicated the space group to be $P2_1$, and hence there are 2 independent molecules in the asymmetric unit and an unknown number of water molecules.

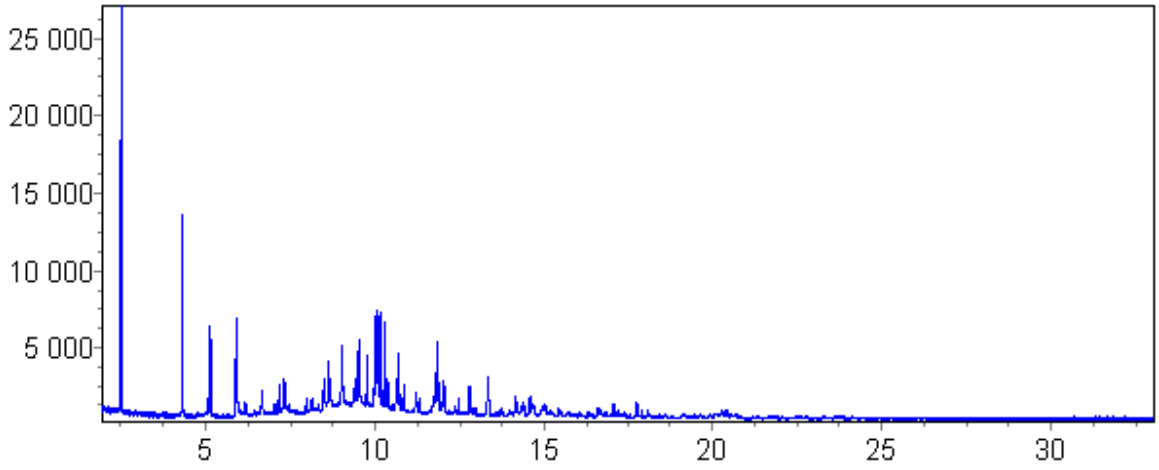


Figure 4-11Diffractogram of the Leu-enkephalin obtained form a solution formed by 10 mg of sample and 1 ml of water measured on ID31.

This phase is not reported in any work in the literature. We tried to solve the structure from the powder data using FOX and TOPAS but with no success. We then tried to solve the structure form the single crystal data using direct methods and the program SHELXS but with no success. A cause of the failure is the disorder of the sample. In these images collected on ID11 we can see a few stripes of reflections that indicate that along a direction in the crystal packing there is disorder.

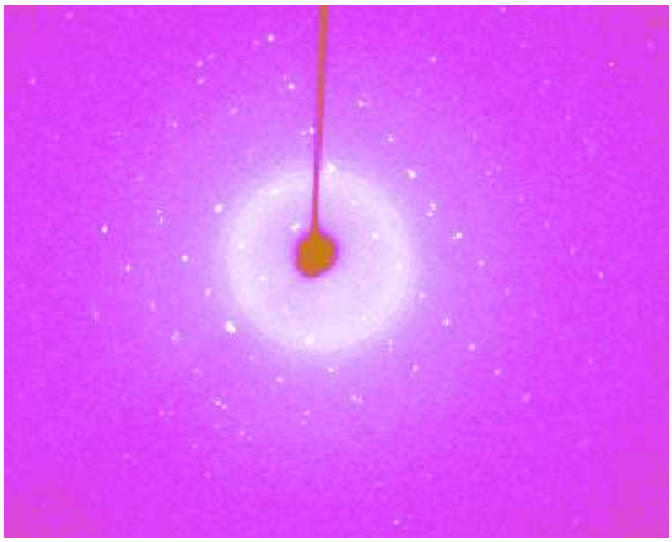


Figure 4-12 Images collected on ID11ot the micro-crystal obtained form a solution formed by 10 mg of sample and 1 ml of water

We then tried to crystallize the sample changing the concentration and the amount of the solution in order to slow down the evaporation process. We let evaporate in the same conditions in the desiccator a solution formed by 30 mg of sample and 4 ml of water. This time we obtained some nice crystals that we measured on ID11.

The structure was solved using direct methods with SHELXS and refined with SHELXL. We don't report here the details of the refinement because the structure that we've found is the same as the one from the work of Griffin et al[65]. The sample crystallize in a monoclinic cell

$$a = 11.647(8)\ \text{Å} \quad b = 15.64(1)\ \text{Å} \quad c = 16.72(1)\ \text{Å} \quad \beta = 92.16(2)^\circ$$

space group $P2_1$, 4 leu-enkephalin and 2 water molecules in the unit cell. The conformation of the leu-enkephalin is extended anti-parallel β-sheet.

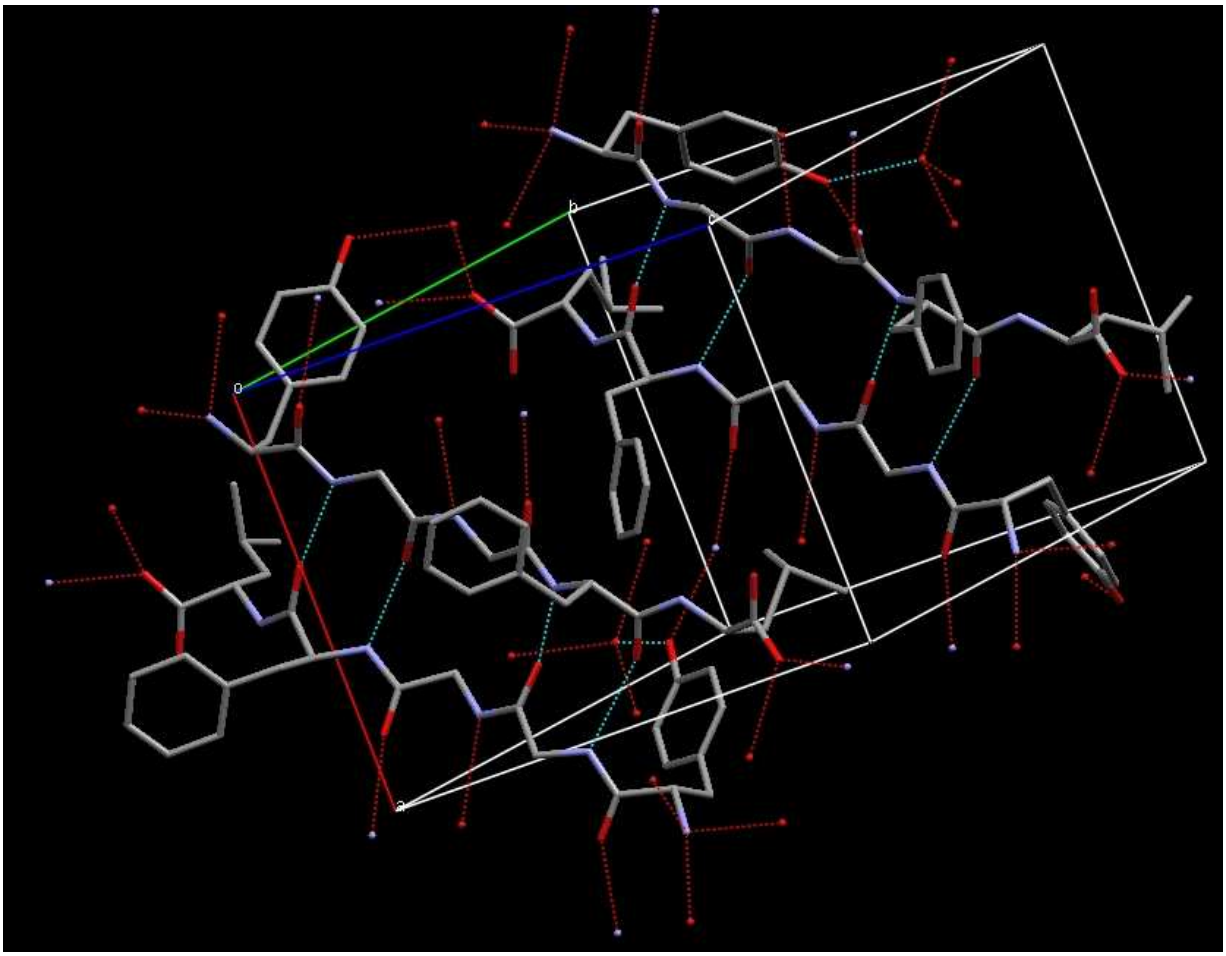


Figure 4-13 Crystal structure of the Leu-enkephalin obtained from solution formed by 30 mg of sample and 4 ml of water and measured on ID11. The conformation is extended anti-parallel β-sheet

We would like to perform a measurement of this powder on ID31 to see at least if it is possible to solve such a complex structure only formed of light atoms with simulations. But as this one was the last sample analyzed there was no time for further measurement.

It could be interesting to try different solution concentrations to see if it is possible to obtain an ordered sample of the phase not solved.

Chapter 5 Localization of Benzene in a Sodium-X Zeolite

Dans ce chapitre le sujet d'étude est la localisation d'une molécule de benzène adsorbée sur des zéolithes X au sodium de structure déjà connue. Ce résultat a été obtenu par simulation et l'utilisation de cartes d'entropie maximale.

Since their first successful use as catalysts between the years 1959 and 1962, synthetic faujasites have been used as fundamental catalytic components for many important processes in the hydrocarbon transformation area[72]. The catalytic properties of these materials are influenced by many factors, including the nature of the exchangeable cations, the aluminum content, and the activation conditions. In addition, the behavior of the adsorbed hydrocarbons plays an important role in catalytic processes that take place within the zeolite cavities. For this reason, the interpretation of the catalytic behavior requires a knowledge of not only the cation distribution within the zeolite cavities but also of the interactions between the cations and the adsorbed molecules.

The framework is formed by SiO_4 and AlO_4^- tetrahedra the global negative charge of which is balanced by exchangeable cations (M^{q+}) distributed over several sites throughout the structure. The chemical composition of faujasites is $M^{q+}_{x/q}Al_x\ Si_{192-x}O_{384}$ with $0<x<96$.

If Si/Al>1.5 then the structure is of "Y-type" and the SiO_4 and AlO_4^- tetrahedra are (long-range) disordered, with space group Fd3m. If Si/Al<1.5 then the structure is of "X-type" and the SiO_4 and AlO_4^- tetrahedra are ordered, space group Fd3.

Figure 5-1 shows a schematic representation of the extra-framework cation positions in zeolites X and Y.

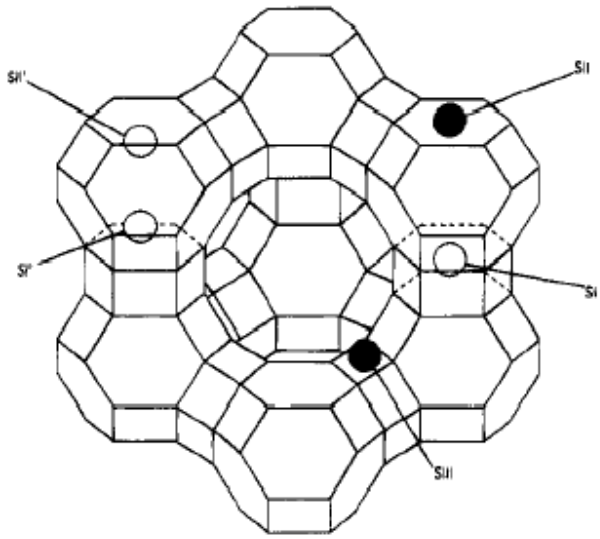


Figure 5-1 Schematic representation of the extra-framework cation positions in zeolites X and Y

Cation distributions in faujasite-type (FAU-type) zeolites have been widely studied by X-ray and neutron diffraction[73-78].

Two important adsorption sites for benzene adsorbed in zeolite Na-Y have been reported in a low-temperature powder neutron diffraction study[79]. The first, which is thought to be occupied predominantly at low coverages, e.g. one benzene molecule per supercage, is found in the supercage, where it is facially coordinated to the site II sodium ions. The second benzene site is located in the 12-ring windows that link together neighboring supercages; this second adsorption site becomes important at higher concentrations of benzene. At low coverage the site II dominates, whereas at higher coverages there is a better balanced between the two possible sites, but with site II still more highly occupied.

In zeolite Na-X instead according to the single crystal X-ray diffraction study by Aubert et al.[80] no evidence for absorption on the site II is seen. They could only observe displacement of cations III, but the benzene wasn't localized anywhere.

The structure of Na-X zeolite, $Na_{86}Al_{86}Si_{106}O_{384}$, containing chemisorbed deuterated benzene has been then studied for two different coverages at 1.5K and room temperature by powder

neutron diffraction. Measurements were performed with the D2B diffractometer at the ILL neutron source in Grenoble.

5.1 Bare Na-X

The first measured sample was the bare zeolite in order to confirm the position of the Na cations. The coordinates for the starting model were taken from the single crystal work of Olson[76] both for the framework and for the cations. The structure was refined with GSAS. Initially we refined the framework and the cations positions and their occupancy then for the cations the occupancy was fixed and we refined the thermal parameters. We then preformed different refinements where alternatively we refined either the occupancy and the positions of the cations or the atomic displacement parameters and positions of the cations. As atomic displacement parameters and occupancy are higly correlated in the final refinement we decided to fix the occupancies at the sensible values that we had obtained and refined just the thermal factors.

The structure is cubic with a= 24.9276(4) Å, the space group is Fd3. The final refinement converged to R_{wp} = 3.4 %, R_p= 2.6 % and χ^2=3.8.

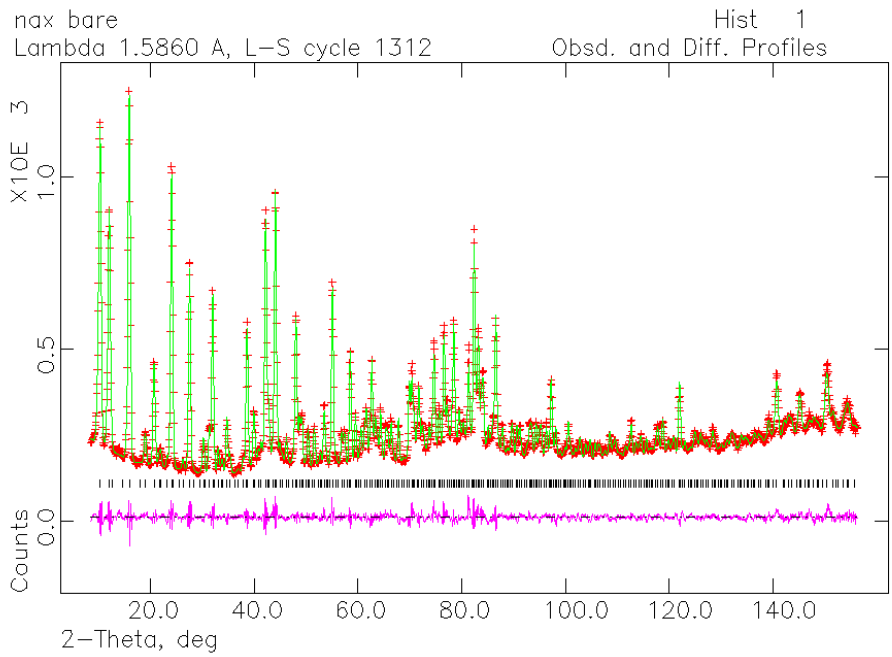


Figure 5-2 Observed (+,red), calculated(-, green) and difference plot of the final Rietveld refinement of the bare Na-X zeolite

In the following tables are reported the final refined framework coordinates:

atom	x	y	z	occ
Al	-0.05412(23)	0.12527(29)	0.03647(31)	0.896
Si1	-0.05412(23)	0.12527(29)	0.03647(31)	0.114
Si	-0.05519(20)	0.03508(30)	0.12450(28)	1.000
O1	-0.10901(23)	0.00399(29)	0.10564(24)	1.000
O2	-0.00264(19)	0.00002(19)	0.14126(14)	1.000
O3	-0.03440(15)	0.07309(23)	0.07519(23)	1.000
O4	-0.06897(17)	0.06992(27)	0.17700(26)	1.000

And in the next ones we report distance and angle values

Bond length (Å)	Si-tetrahedra	Al/Si-tetrahedra
T-O1	1,621(5)	1,682(5)
T-O2	1,632(5)	1,704(8)
T-O3	1,631(5)	1,692(5)
T-O4	1,611(5)	1,679(5)
average	1,623(6)	1,689(6)

Angle (°)	Si-tetrahedra	Al/Si-tetrahedra
O1-T-O2	117,5(5)	109,6(5)
O1-T-O3	108,1(4)	112,5(4)
O1-T-O4	108,9(5)	110,2(5)
O2-T-O3	104,6(4)	106,6(4)
O2-T-O4	105,7(5)	106,3(4)
O3-T-O4	112,1(5)	111,4(5)
average	109,4(5)	109,4(5)

Angle (°)	
Si-O1-Al	135,2(3)
Si-O2-Al	145,2(3)
Si-O3-Al	141,9(4)
Si-O4-Al	146,3(3)
average	142,1(3)

5.1.1 Cations location

We located the cations in the following positions:

Site I cations

Na1(0,0,0), occupancy of 0.042 , the site has multiplicity 16 that means that the site has occupancy for unit cell of 0.672.

Na1(0,0,0)	distance(Å)
Na1-Al	3,503(7)
Na1-O3	2,754(4)

Site I' cations

There are two cations Na in two I' postions.

Na2, (x, x, x) x=0.0457(5), occupancy of 0.689, the site has multiplicity 32 that means that the site has occupancy for unit cell of 22.048

Na2 (0,0457(5), = ,=)	distance(Å)
Na2-Al	3,184(7)
Na2-O2	2,903(4)
Na2-O3	2,243(5)

Na2', x=0.0530(12), occupancy of 0.311, the site has multiplicity 32 that means that the site has occupancy for unit cell of 9.952

Na2' (0,0530(12),=,=)	distance(Å)
Na2'-Al	3,245(15)
Na2'-O2	2,919(8)
Na2'-O3	2,313(17)

When we first try to refine one single cation in the I' position, we ended up with a too big final thermal parameter for the Na2 atoms.

The site I' is in the center of a 6-ring, in the unit cell there are 32 rings of this kind. The number of Al atoms is 86. This means that on average there are 22 six-rings where 3 tetrahedral sites are substituted with aluminum and in the other 10 six-rings there are 2 aluminum. This explains the presence of two different cations in the position I'.

The longer Na-O distance for cations Na2' is consistent with the lower Al contents for these 10 six-rings.

Site II cations

There are two Na II positions, for the same reason explained for site I'.

Na3 (x, x, x) x=0.2284(3), occupancy of 0.689, the site has multiplicity 32 that means that the site has occupancy for unit cell of 22.048

Na3 (0.2284(3), =, =)	distance(Å)
Na3-Al	3,223
Na3-O2	2,320
Na3-O4	2,862

Na3' (x, x, x) x=0.2388(7), occupancy of 0.311, the site has multiplicity 32 that means that the site has occupancy for unit cell of 9.952

Na3' (0,2388(7), =, =)	distance(Å)
Na3'-Al	3,336
Na3'-O2	2,472
Na3'-O4	2,917

The occupancy value is the same as for cation site I'. In fact as the thermal parameters are very correlated with the occupancy in a powder refinement it is very difficult to determine the exact value. From the first refinements we saw that the value obtained was near the expected value having 2.2 times more density on the site Na2,3 than Na2,3'. We then fixed them at the final values reported for both site I' and II.

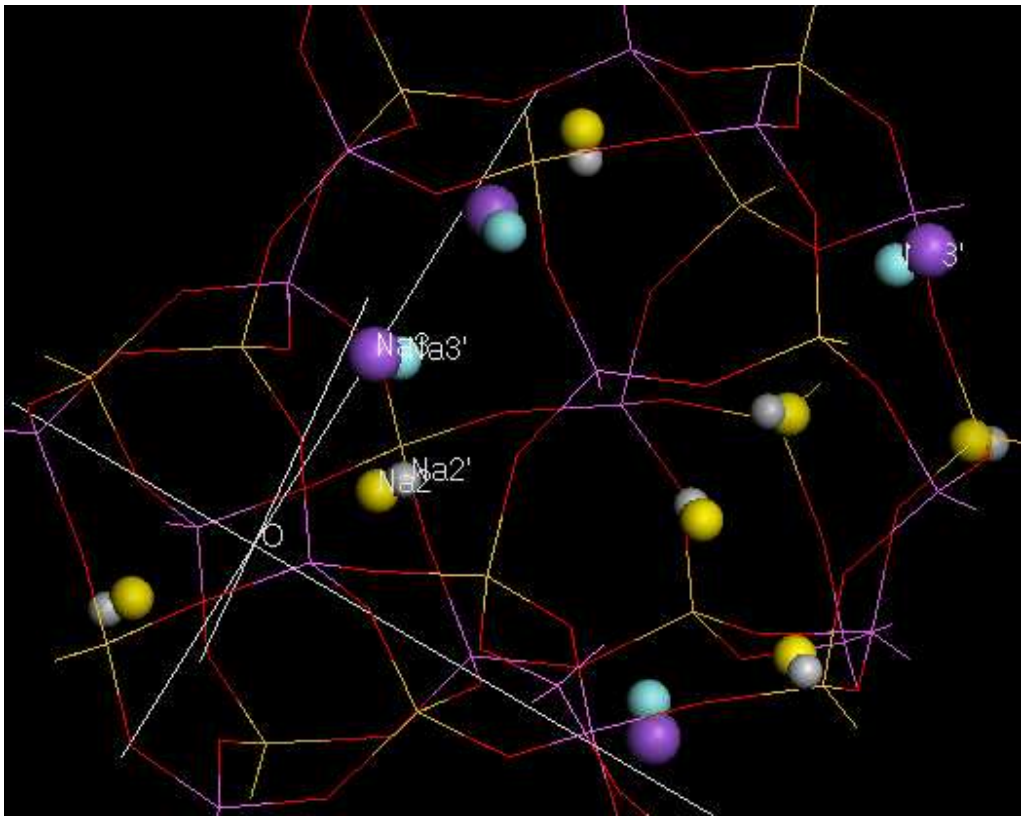


Figure 5-3 Location of cations in site I' and II.

Site III' cations

Two positions where found near the idealized III site.

Na6 (0.363(2), 0.352(2), 0.1415(2)), occupancy of 0.1027, the site has multiplicity 96 that means that the site has occupancy for unit cell of 9.859

Na6 (0.363(2), 0.352(2), 0.1415(2))	distance(Å)
Na6-Al	2,75(6)
Na6-O1	2,85(5)
Na6-O3	2,92(5)
Na6-O4	2,56(6)
Na6-Na7	2,6(9)

Na7 (0.438(2), 0.282(2), 0.161(2)), occupancy of 0.1034, the site has multiplicity 96 that means that the site has occupancy for unit cell of 9.931

Na7 (0.438(2), 0.282(2), 0.161(2))	distance(Å)
Na7-Al	3,15(5)
Na7-O1	2,59(6)
Na7-O4	2,51(5)

The sum of the total Na occupancy for the unit cell is 84,46 very near to the expected value of 86.

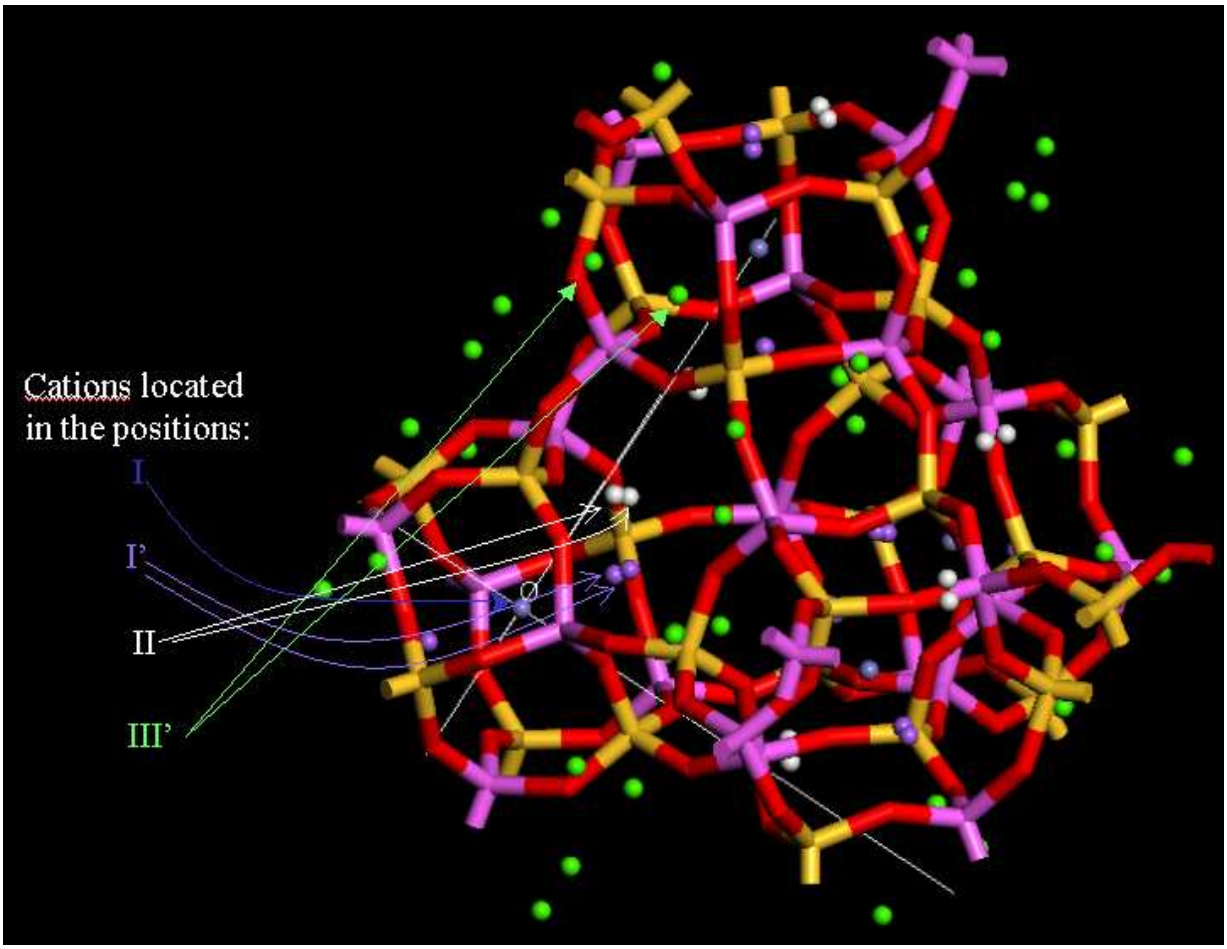


Figure 5-4 Localisation of Na atoms in the structure

5.2 Benzene location

Using the model that we just described we tried then to fit the data of the sample with the chemisorbed benzene, and keeping everything fixed. Figure 5-5 shows the fit of the sample with 1 benzene chemisorbed per supercage at room temperature.

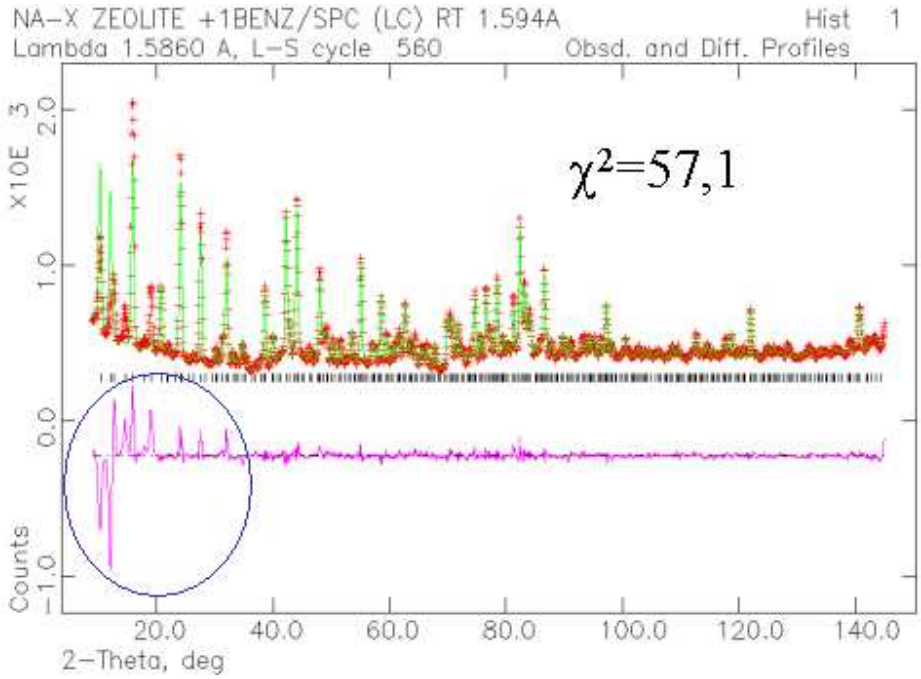


Figure 5-5 Observed (+,red), calculated(-, green) and difference plot of the Rietveld refinement of the Na-X zeolite with 1 benzene chemisorbed per supercage, excluding the benzne in the model. The fit is very bad

The fit is very bad especially in the low angle region. This is to be expected because we are fitting data from the sample with benzene chemisorbed with the model of the bare zeolite where the benzene is not present. We tried as a first attempt to use the model of the Na-Y zeolite with chemisorbed benzene but the fit didn't improve; there is no evidence for the location of the benzene facing site II. We then tried to perform a difference Fourier map using the F_{obs} extracted from this fit and the result is shown in Figure 5-6:

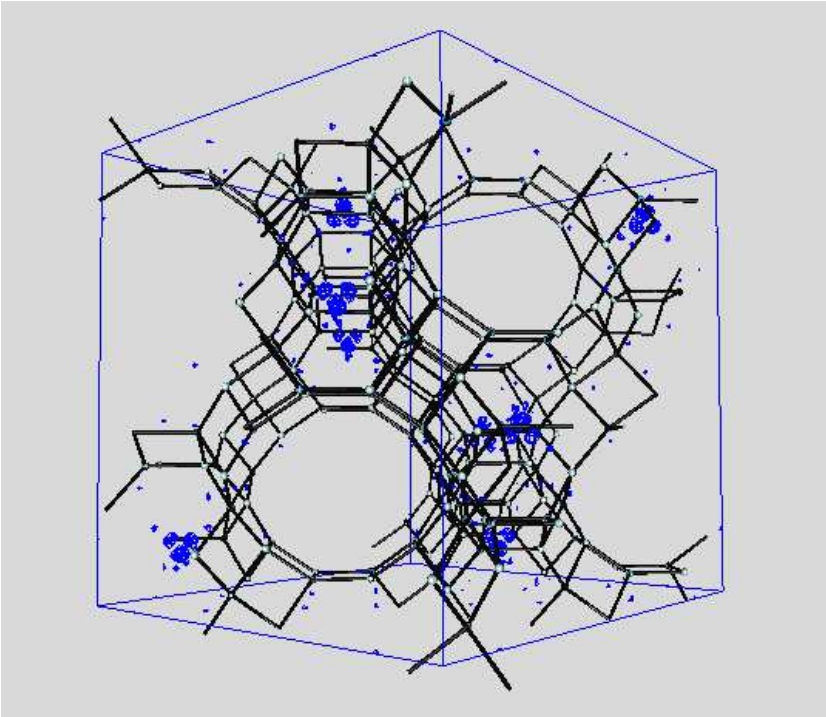


Figure 5-6 Difference Fourier map using the Fobs extracted from the fit of Figure 5-5

with no evidence for the benzene. Some residual density is only appearing inside the framework but there is no good reason to find any density over there: it is not possible for the benzene to go inside the sodalite cage as benzene is too big.

Then from the same F_{obs} we computed a MEM map starting from a uniform prior, with the program PRIMA.. The result is shown in Figure 5-7.

Figure 5-7 MEM map starting form a uniform prior

It is possible now to recognize extra density in the zeolite channel and this is where, as we will see, the disordered benzene is found.

For comparison this is the result that could be obtained using exactly the same Fobs but performing a Fourier synthesis:

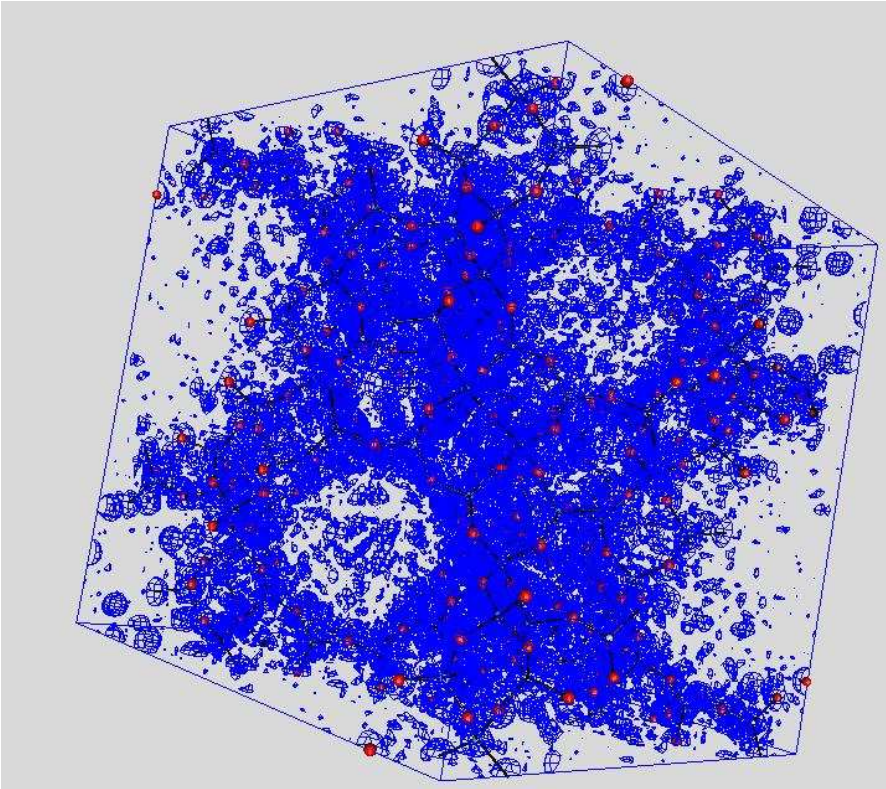


Figure 5-8 Fourier map using the same Fobs

It is not possible to locate the benzene as already seen form the difference map in Figure 5-6 Even from the MEM map it is not possible to locate more than the general whereabouts of the benzene. Therefore we tried to perform a simulation. Using the program FOX and the parallel tempering algorithm we fixed the structure of the bare zeolite and we inserted a benzene molecule as a rigid body. The orientational and rotational parameters of the rigid body were left free to vary as well as the occupancy since the nominal value of 1 molecule par supercage might in reality be slightly different. It is then an optimisation with 7 free parameters. The solution was quickly found and the benzene is located on the 12-ring windows, interacting with cations III'. Those cations are displaced after the introduction of the benzene as seen in the work of Auber et al[80]. For each different benzene coverage, we repeated the simulation from the beginning and we always found a similar result.

5.2.1 Final refinements

The first sample studied was the Na-X zeolite with one benzene chemisorbed per supercage at room temperature.

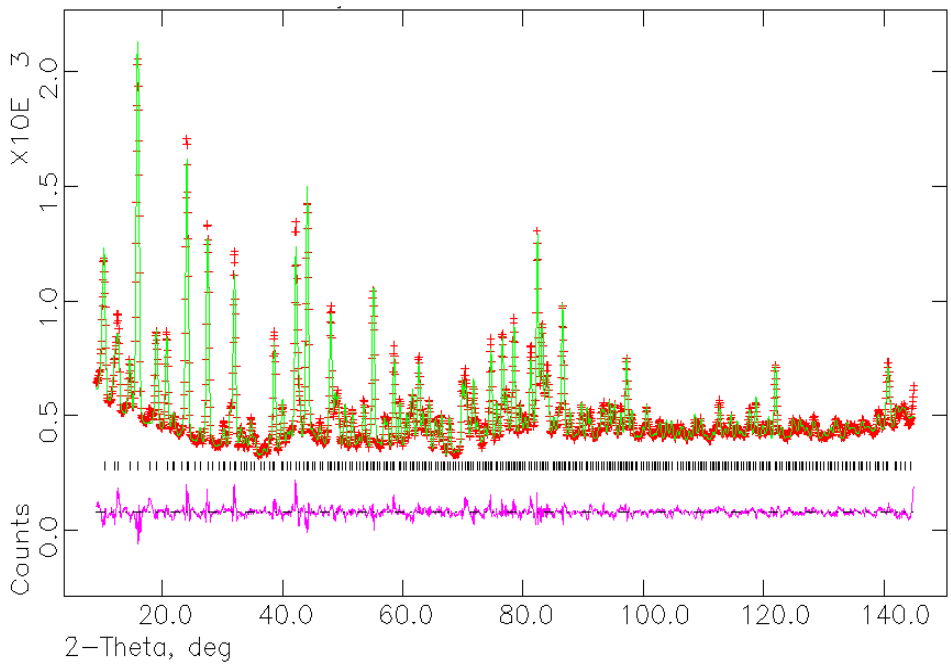


Figure 5-9 Observed (+,red), calculated(-, green) and difference plot of the Rietveld refinement of the Na-X zeolite with 1.44(7) benzene molecules per supercage

The structure is cubic with a= 24.9317(4) Å, the space group is Fd3. The final refinement converged to R_{wp} = 3.4 %, R_p= 2.6 % and χ^2=9.7. The amount of benzene coming from the final refinement was actually 1.44(1) benzene molecules per supercage. This means that there is an average of 44% of supercages having 2 benzenes chemisorbed and 56% having a single one, supposing that the benzene is homogenously distributed in the structure.

In Figure 5-10 and Figure 5-11 the location of the disordered benzene is shown

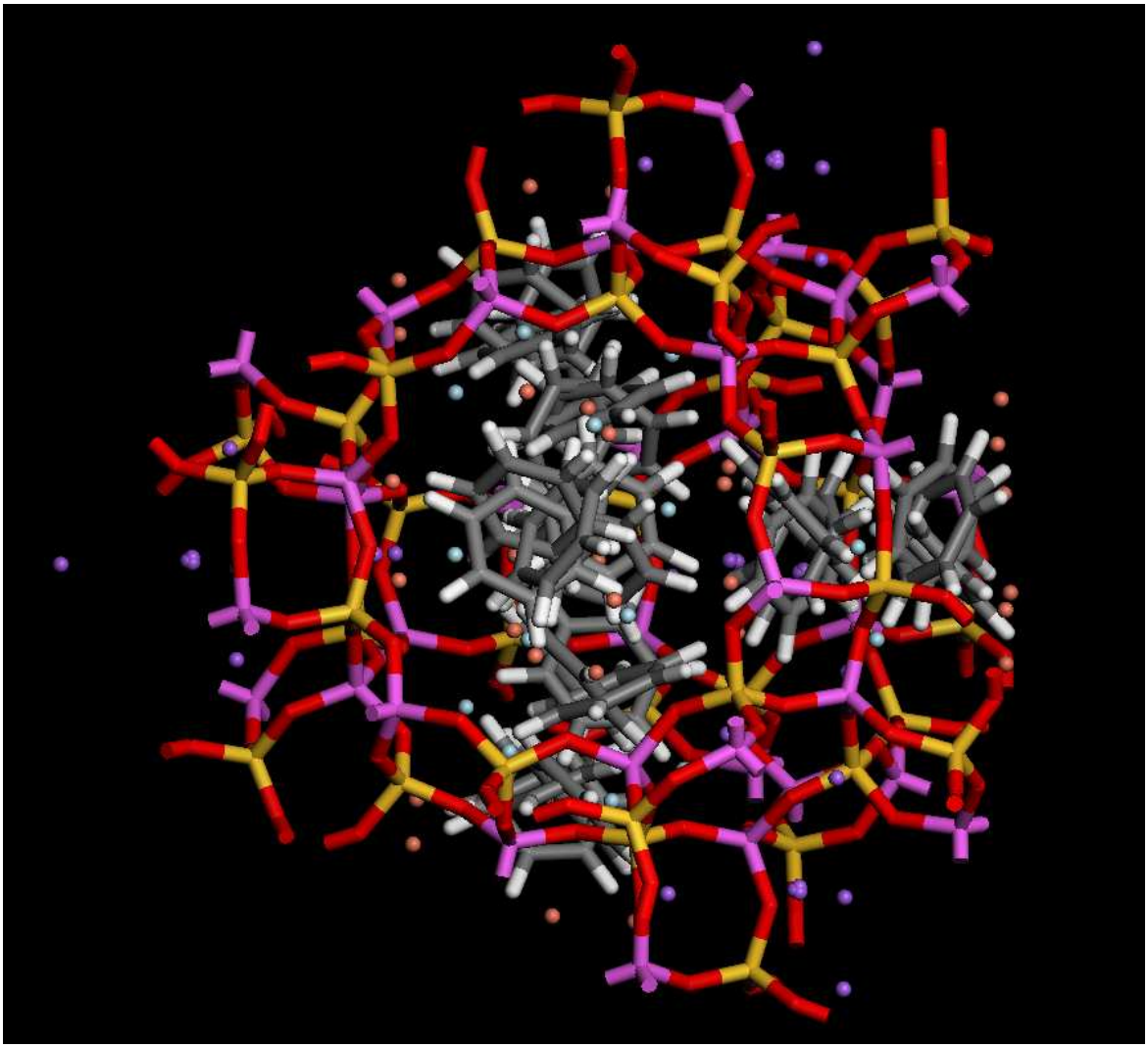


Figure 5-10 Benzne location on the on the 12-ring windows

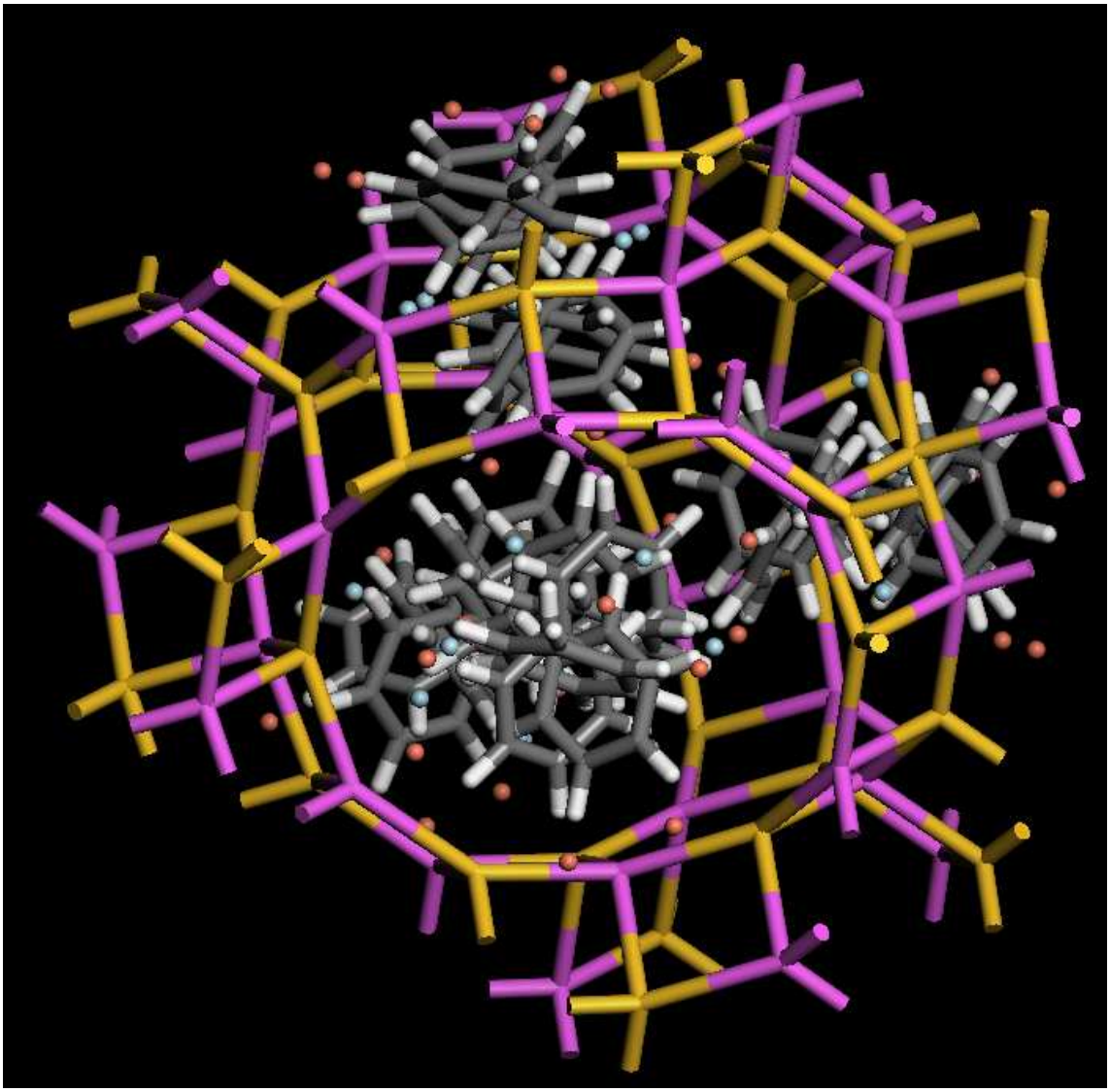


Figure 5-11 Benzene location on the 12-ring windows, the oxygen have been omitted in order to clarify the picture and also all the Na atoms except Na6 and Na7.

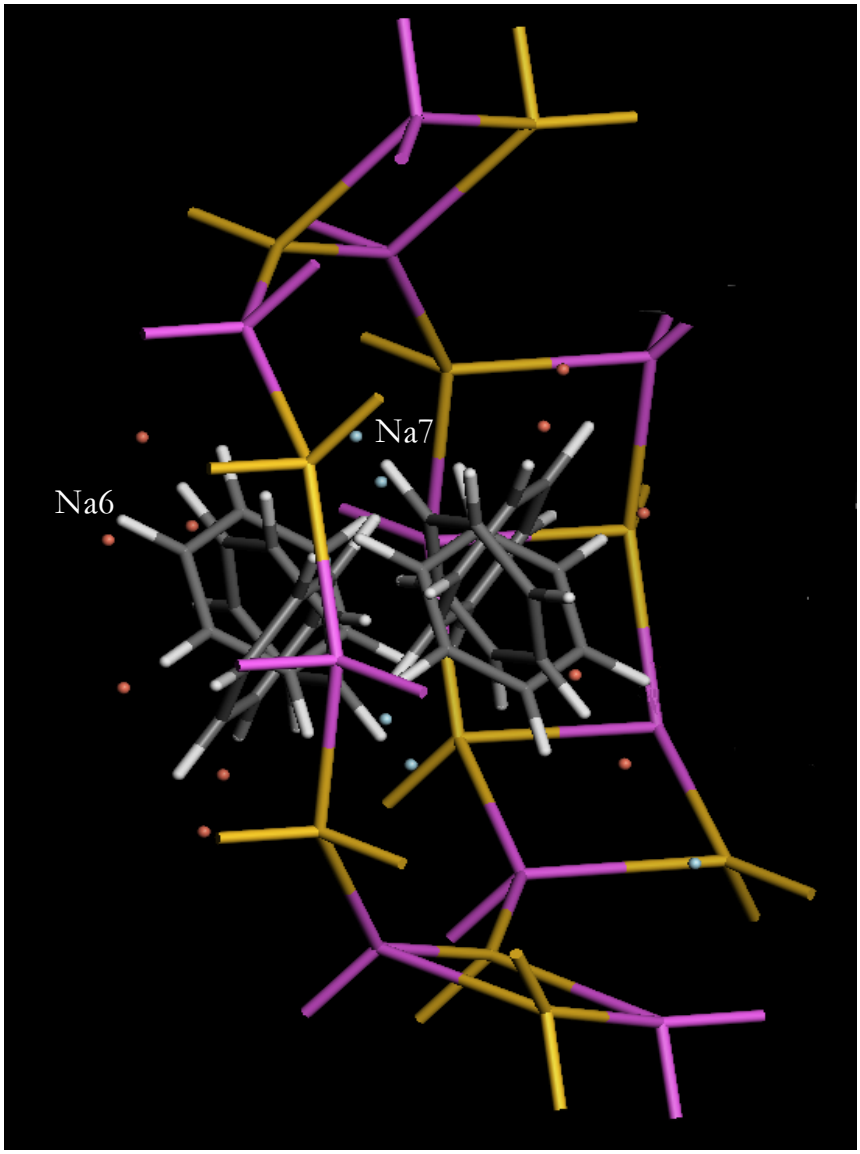


Figure 5-12 Zoom of Figure 5-11

Cations Na6 are displaced after the introduction of the benzene. They interact less with the framework and they also interact with the benzene.

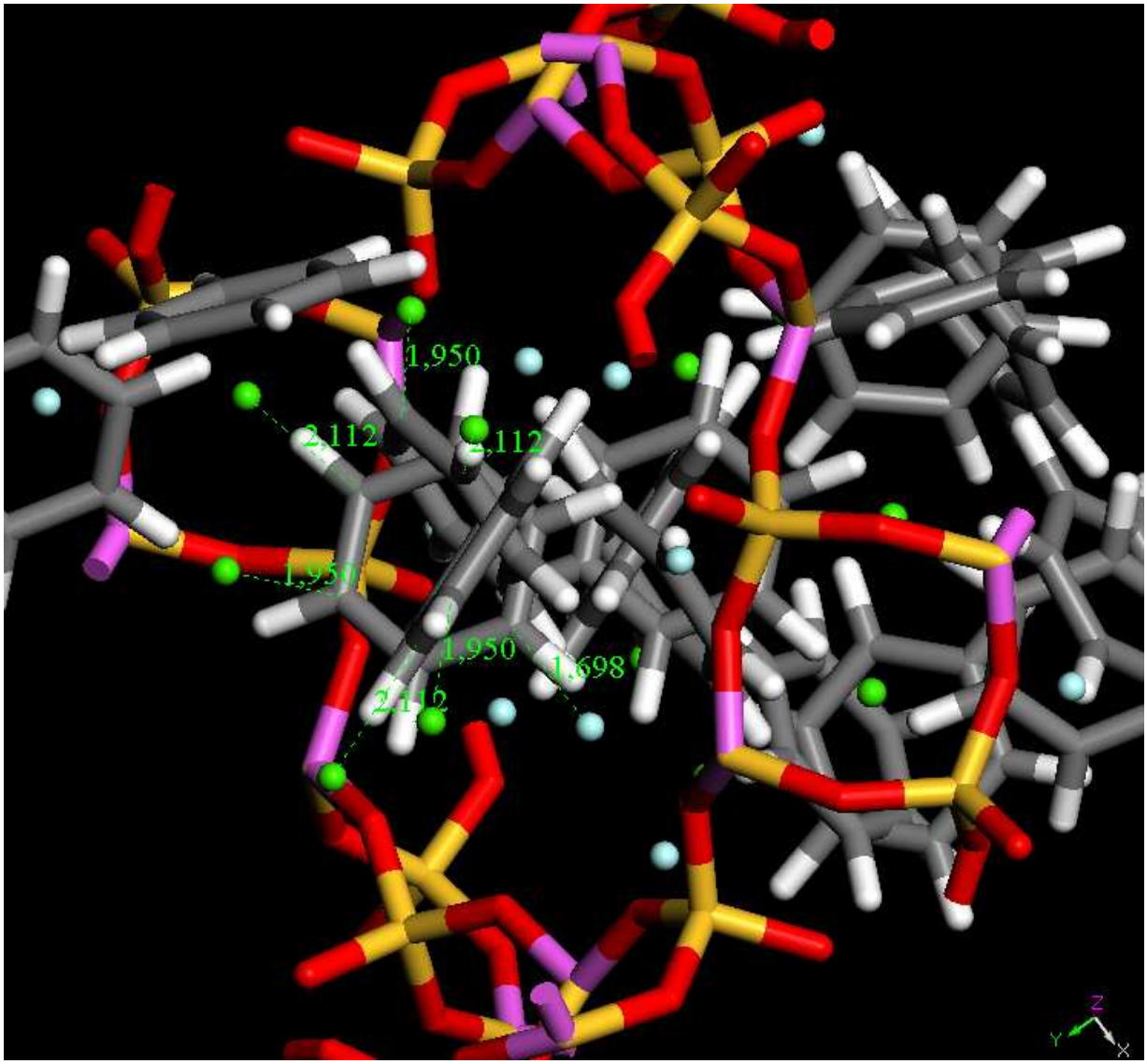


Figure 5-13 Interaction of cation Na6 with the benzene

The cations in position III' are situated on the 12-ring windows, the Na7 in blue are in the plane of the window (Na7 sample without benzene, Na7b sample with benzene), in green the Na6 that are facing the supercage (Na6 sample without benzene, Na6b sample with benzene)

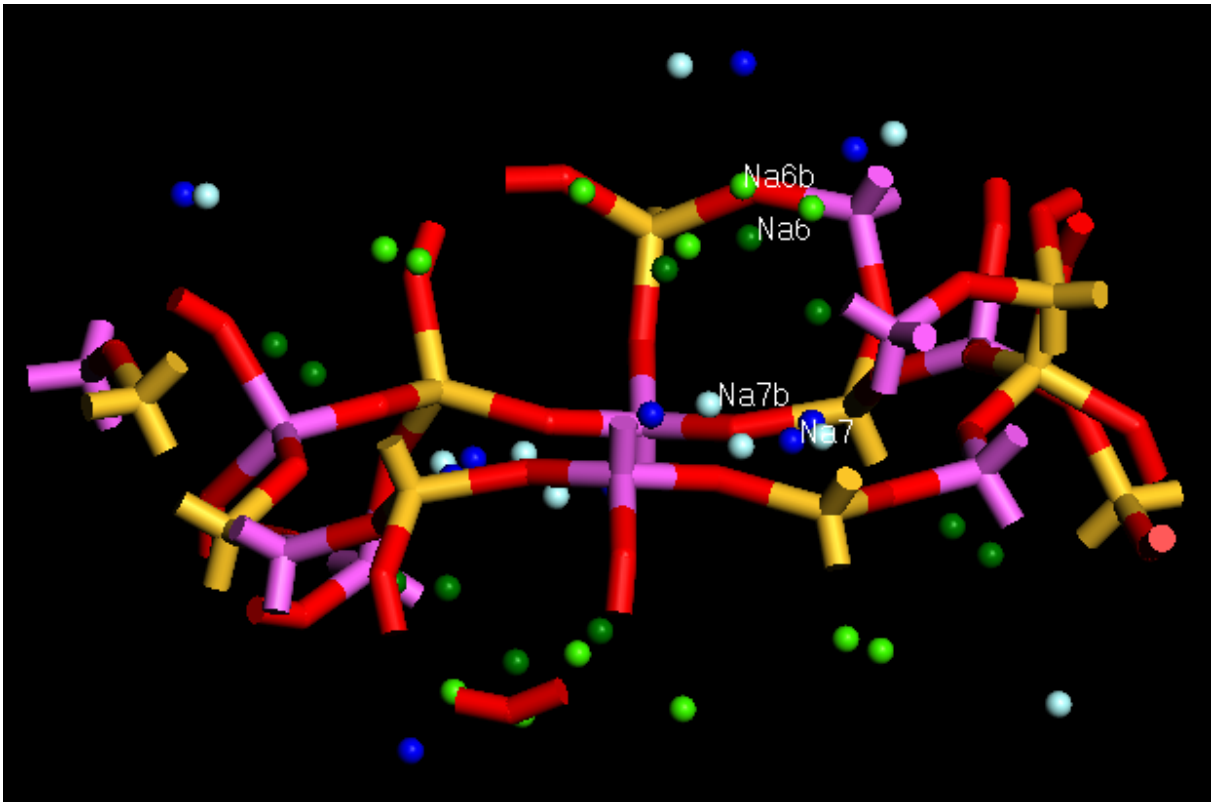


Figure 5-14 Displacement of cations Na6 and NA7 after the introduction of the benzene

As can be see in the following figures, the Na6 cations are largely displaced after the introduction of the benzene and they interact with the benzene. The Na7 cations instead are displaced but not so markedly.

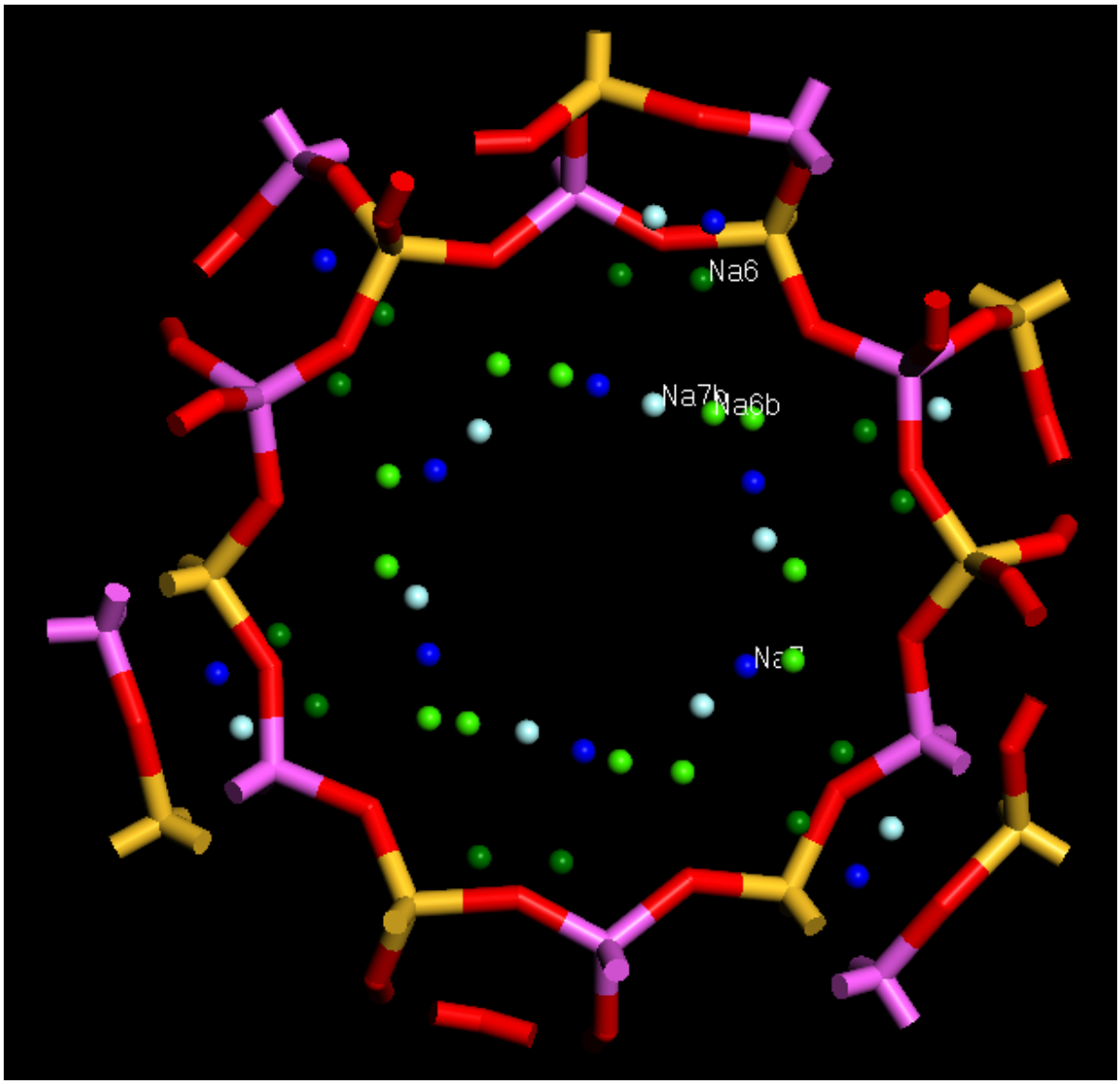


Figure 5-15 Displacement of cations Na6 and Na7 after the introduction of the benzene, another view

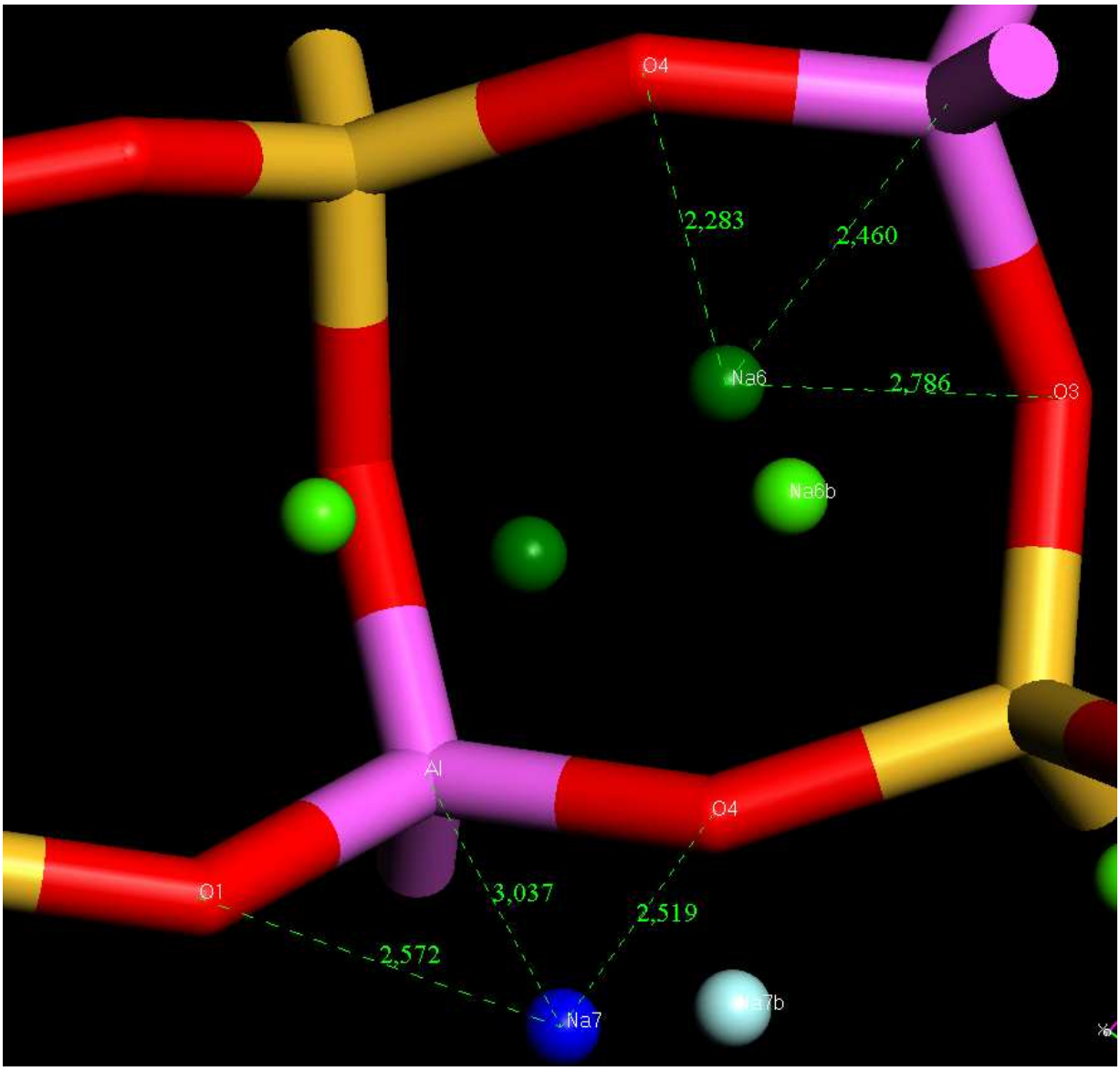


Figure 5-16 Location of cations in the site III' before and after (b) the introduction of the benzene

.

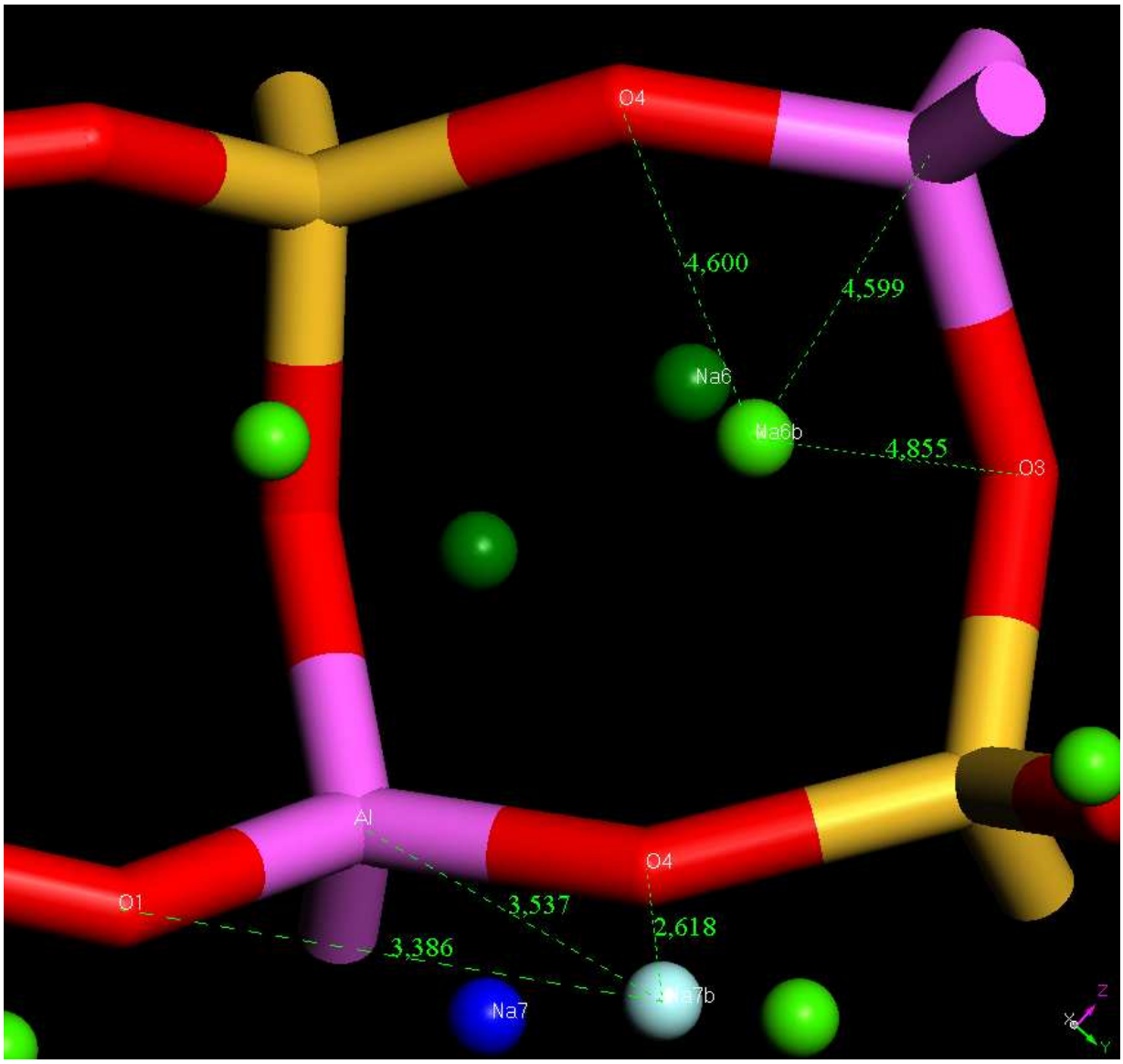


Figure 5-17 Location of cations in the site III' before and after (b) the introduction of the benzene

The second sample studied was the Na-X zeolite with three benzene chemisorbed per supercage at room and low temperature.

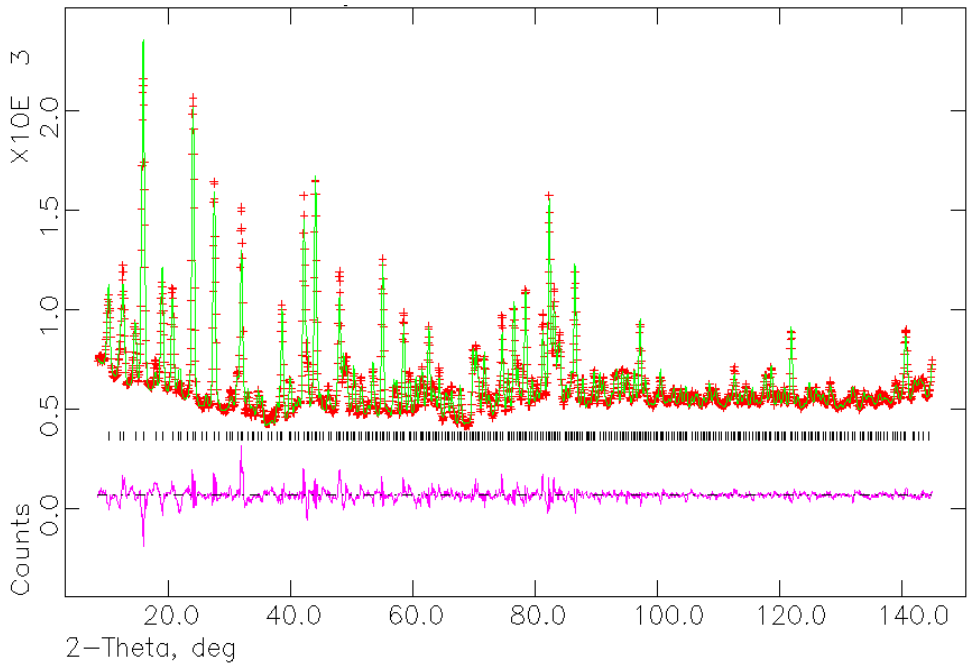


Figure 5-18 Observed (+,red), calculated(-, green) and difference plot of the Rietveld refinement of the Na-X zeolite with 2.07(1) benzene molecules per supercage at room temperature

The structure at room temperature is cubic with a= 24.9310(2) Å, the space group is Fd3. The final refinement converged to R_{wp} = 3.7 %, R_p= 2.7 % and χ^2=9.4. The amount of benzene coming from the final refinement was actually 2.07(1) benzene molecules per supercage. We tried also to perform a refinement where we fixed the occupancy of the benzene to 2.88 molecules per supercage. This value is the one coming form the low temperature measurement, as we will show in the next paragraph. But in this case the thermal factor reached a huge value of U_{iso}= 0.71(1) Å^2. At room temperature the benzene is more disordered and it is very difficult to find the right occupancy value, as it is highly correlated with the atomic displacement parameters. As the benzene is disordered we can actually see less benzene comparing to the low temperature case. The position of the benzene molecules and the cations in the position III' are very similar to the previous sample.

At 1.5 K the structure is cubic with a= 24.9512(4) Å, the space group is Fd3.

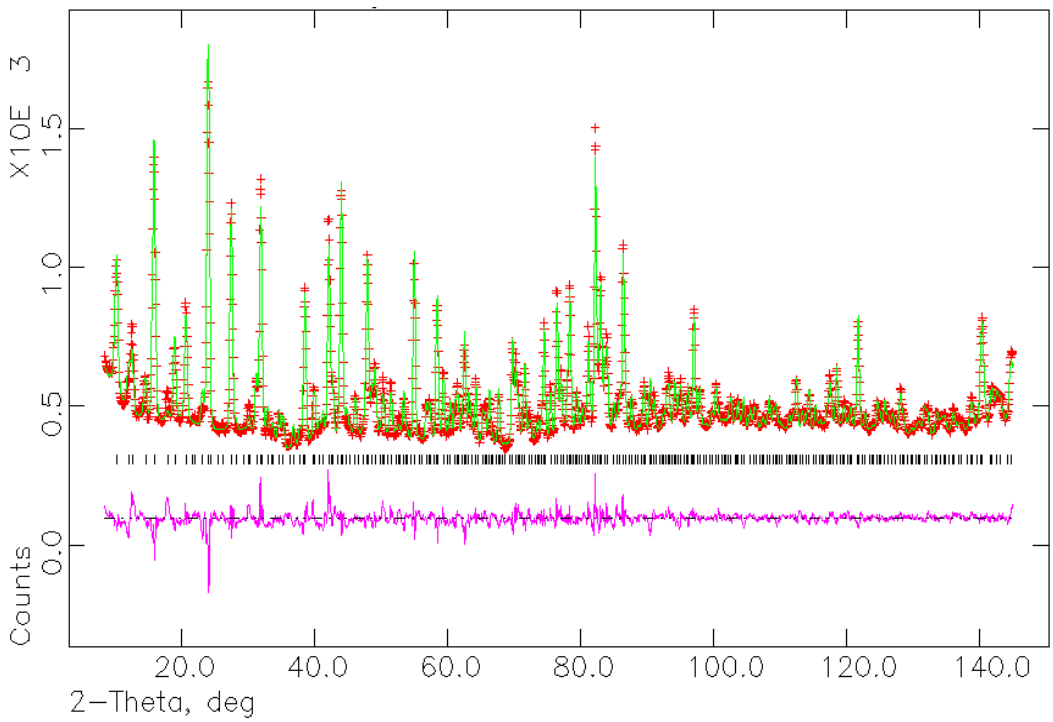


Figure 5-19 Observed (+,red), calculated(-, green) and difference plot of the Rietveld refinement of the Na-X zeolite with 2.88(3) benzene molecules per supercage at 1.5 K.

The final refinement converged to R_{wp} = 3.8 %, R_p= 2.8 % and χ^2=9.1.

For this sample it was not possible to discern the two different positions for the cations in the site I' and II. The amount of benzene coming from the final refinement was actually 2.88(3) benzene molecules per supercage. The position of the cations in the position III' are very similar to the previous sample. In this case further attention must be paid to the benzene position. In fact an amount of 2.88(3) benzene per supercage implies, unlike the previous case, that two of the six benzene positions in Figure 5-20 are occupied at the same time.

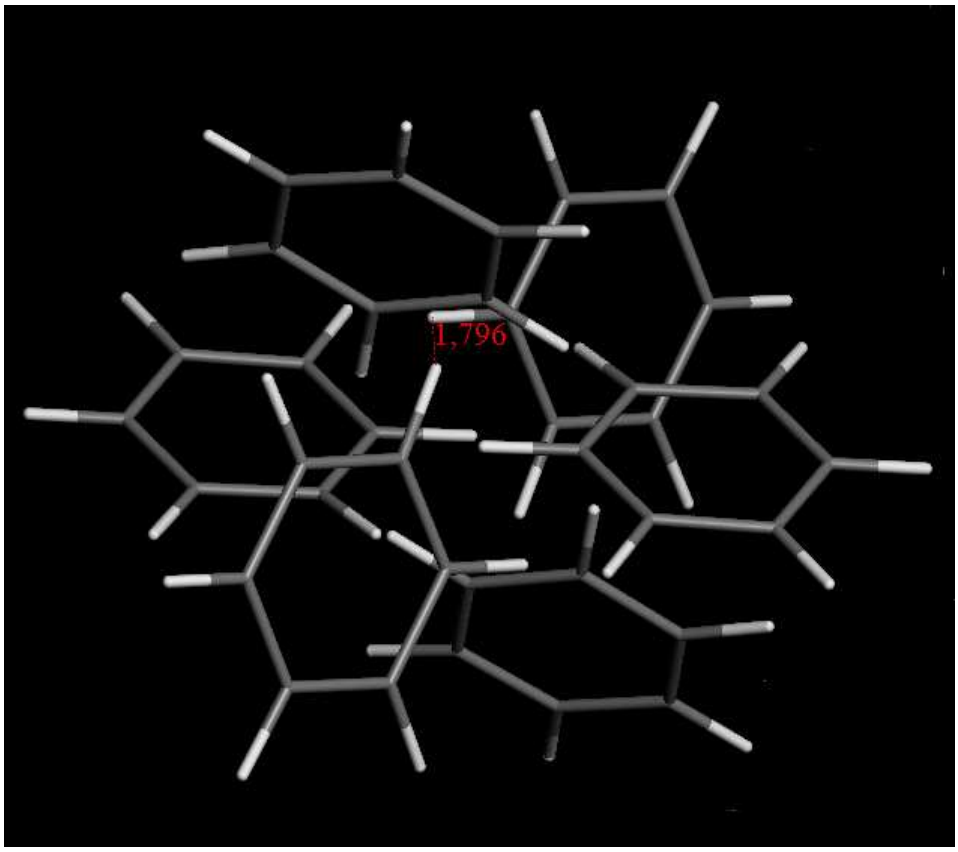


Figure 5-20 Disposition of benzene

This is possible as the minimum distance between 2 benzenes is 1.796(13) Å and this allows the two sites to be occupied at the same time. Comparing to Figure 5-12 it can be seen that now the benzenes are almost in the same position, as the previous lower coverage cases, but more spread out in space.

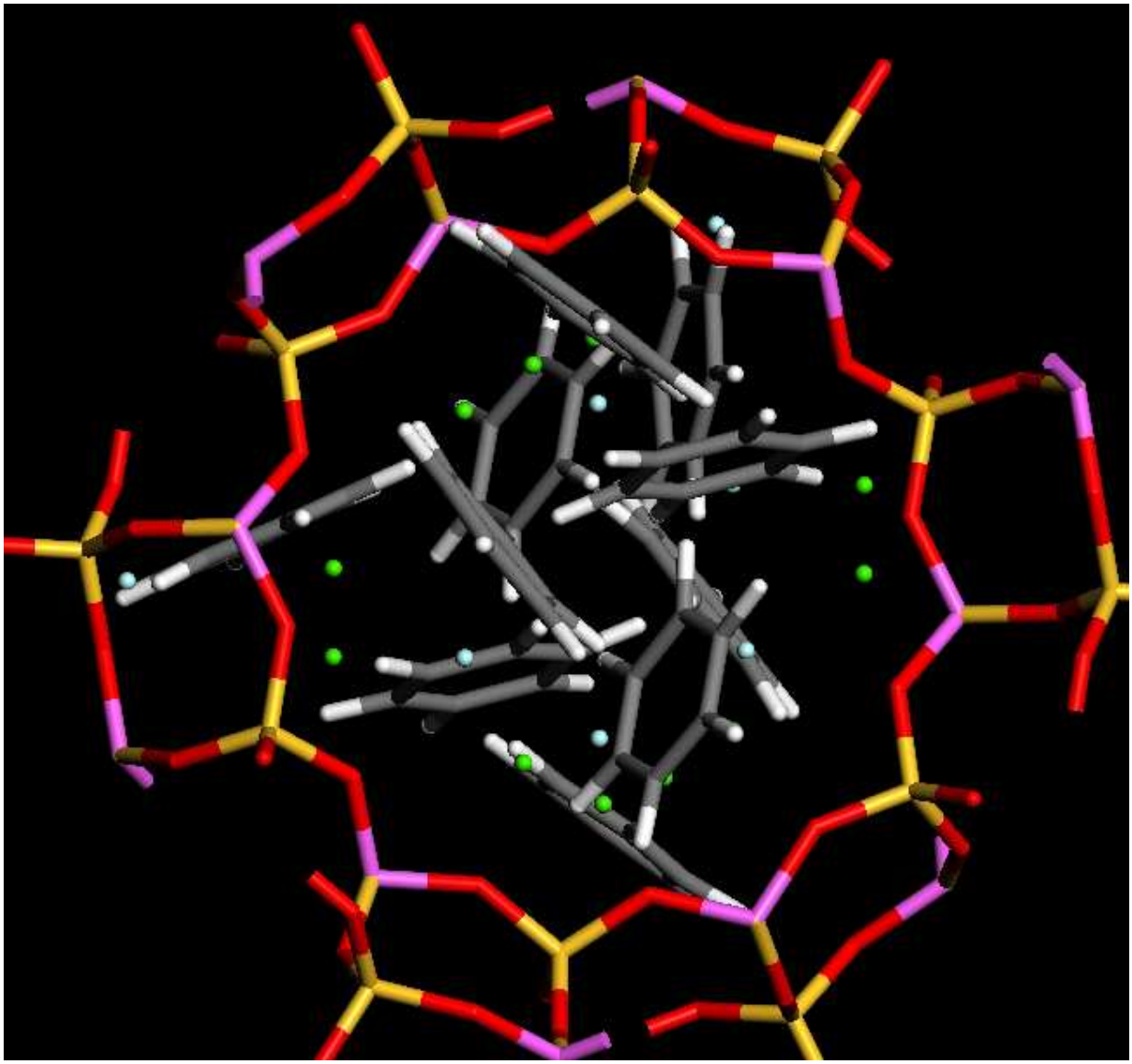


Figure 5-21 Benzene location on the on the 12-ring windows, in order to clarify the picture all the Na atoms have been omitted except Na6 and Na7

To note also is that the unit cell is bigger due to the increased amount of benzene chemisorbed and the repulsion between benzene molecules. In the appendix are reported all the final refined distance and angle values.

Chapter 6 Conclusion

It has been shown in this thesis that X-ray powder diffraction is a powerful tool in crystal structure analysis.

After a brief crystallographic introduction we presented the main peculiarity of a powder diffraction experiments and we emphasized the difficulties due to peak overlap. We described in detail the real space search Monte Carlo algorithms that we used to solve the crystal structures and the way in which the maximum entropy maps are computed showing the mathematical steps and, we also explained the general entropy concept.

We have shown in chapter 3 that we were able to solve several terpene structures and to study the low temperature behavior of those compounds. We found in the literature some information about 7-oxabicyclo[2.2.1]heptane; studies by Ralph M. Paroli et al.[52] using spectroscopic and differential scanning calorimetric techniques, showing that four different solid phases may occur, with hysteresis in the transition temperatures. No previous attempts had been made to solve the crystal structures of those phases, probably because of the difficulties of growing single crystals and the destructive nature of the phase transitions, and in our work we have been able to identify and solve them. The solid-state behaviour of Norcamphor, $C_7H_{10}O$, was studied by Holderna-Natkaniec et al.[59] using neutron diffraction and inelastic incoherent scattering in the temperature range from 10 K to 290K From the neutron diffraction study they didn't report any cell constant and no attempt to solve the structures was made owing to the poor resolution of the diffractograms. Thanks to the high resolution on ID31 we were able to solve them even if some uncertainties remain for the lowest temperature phase because of the presence of the preferred orientation. The other compounds analysed were completely unknown. Some of the samples could not be solved because they present strong granularity or preferred orientation. For such compounds we also tried to simulate the preferred orientation effect but with no success. Trying to

simulate at the same time the model and an effect that can alter the actual values of the intensity is too complicated because the two factors are too correlated. A Gandolfi type spinner could be a solution for such compounds; the sample should be rotated in the largest number of possible directions during the data collection. A complete recapitulative on the results that we obtained has been presented at the end of the chapter.

We presented then in chapter 4 a study on opioid neuropeptides. The most complex structure solved in this work was the Glycine-Glycine-Phenylalanine, where no heavy atoms are present, and the simulation took a few hours on a Pentium 4, 1.5 GHz. We did not find problems in the simulation when the quality of the sample was good, but at the moment there is no real evidence on how complicated the structure can be to be solved by simulation. Further experiments should be performed on more complicated samples. One in this case should pay attention to the fact that the more complex the structure to analyse the more complex will be the N-dimensional space to explore, and then it will became more difficult to be sure about the obtained answer. There could be many similar local minima very near the true minimum and to find it could be very difficult. Help can arise using penalty functions that will help in finding the right solution but they will slow the computational process.

We have shown, in chapter 5 were we described the localisation of the benzene inside a sodium-X zeolite, that simulation can replace successfully the computation of Fourier maps and the manual location of missing atoms in the structure, especially in the case of disorder that is very difficult to find manually. Even from the single crystal X-ray diffraction study of Aubert et al.[80] they could not localize the benzene. We believe that in difficult cases it is better to perform a simulation than to use Fourier maps. In any case we think that if a density map has to be used in order to complete a structure it is better to compute maximum entropy maps that are of better quality and computed, for the cases that we have shown in this thesis, in less than a minute (Pentium 4, 1.5 GHz). These maps are model biased exactly as the Fourier maps because the set of F_{obs} used to compute the maximum entropy map is the same, but the truncation errors present in Fourier maps are avoided. One should still take care when looking at the finest details as explained at the end of the entropy paragraph and this is still an open field of research.[41, 42]

We have seen that almost all the samples have been indexed owing to the high resolution available on ID31, but any impurity in the sample causes the main problem in indexing the diffractogram: even a small amount of impurity will be detected, owing to the very high statistical quality of the data. Indexing programs can take this into account but even if they can find the solution it is sometimes very difficult to distinguish it from all the possible proposed cells.

From an experimental point of view, we encountered some difficulties, at the beginning, because the organic samples degraded under the X-ray flux in a few minutes. Because of this we had continuously to close the shutter enter in the experimental hutch and translate the capillary a few millimetres in order to take a fresh part of the sample. This was possible at room temperature but it was very tedious and almost impossible when the cryostream was in use. A solution has been recently found: a new cryostream is now in use where the sample can go into the nozzle with no problem and a motorized spinner can automatically translate the sample. A robot has been installed too, and this allows automatically the changing of the capillary when the whole capillary is burned or when one has to process a series of capillaries at night. This is why we couldn't really try, during these last two years, to collect multiple data sets at different temperatures, in order to follow the shifts of the diffracted peaks relative to different reflections, exploiting the anisotropic thermal expansion of the solids. In this manner a sort of "pseudo-single crystal" approach may be applied to the most difficult-to-index powder data patterns. The same principle may be applied for the partitioning of overlapping diffracted intensities to get better-resolved experimental extracted intensities for the determination of more complex structures. But in any case we have shown that the simulation was already good enough to solve all the structures that we analysed, mainly using the two programs: FOX and TOPAS.

In the future it will be very interesting to continue the work on neuropeptides or on other biological samples crystallized without the use of any organic solvent to see differences with the single crystal work or for still unknown structures. We should try to exploit the anisotropic thermal expansion now that most of problems have been fixed on the beamline

ID31 with regard to coping with radiation damage. A further improvement on the beamline could be also the use of an area detector to check the crystallinity of the sample and to be sure that there is no preferred orientation or granularity in the sample that affect the intensity values and complicate a lot the structure solution stage. An alternative development might be a one-dimensional strip detector for collecting data rapidly over a large angular range, for samples that are highly radiation sensitive, or for a detailed investigation of the phase-change behaviour of a sample.

Conclusion(Fr)

Il a été démontré dans ce travail de thèse que la diffraction de rayonnement X sur poudres est un puissant outil d'analyse de la structure cristalline.

Après une brève introduction générale à la cristallographie, nous avons présenté les caractéristiques principales d'une expérience de diffraction X de poudres, à propos de laquelle nous sommes revenus sur les difficultés liées au recouvrement des pics. Nous avons décrit dans le détail les algorithmes de recherche de Monte Carlo dans l'espace direct que nous avons utilisés pour atteindre les structures cristallines. De plus, après avoir exposé les concepts généraux de la fonction d'entropie, nous avons détaillé les étapes successives permettant d'obtenir la carte d'entropie maximale.

Nous avons montré, dans le chapitre 3, que nous sommes capables de résoudre la plupart des terpènes et d'étudier leurs comportements à basse température. Par exemple, M. Paroli et al [51] ont étudié le 7-oxabicyclo[2.2.1]heptane grâce à des expériences de spectroscopie et de calorimétrie. Ils ont montré que 4 phases solides existent, et qu'il y a une hystérésis à la température de transition. Cependant, vu la difficulté à obtenir un monocristal et la nature destructive de la transition de phase, ils n'y a pas eu de tentatives de résolution de la structure de ces phases. Au cours de notre travail, nous avons été capable d'identifier les différentes phases et de les résoudre. Un autre exemple concerne le comportement dans l'état solide du Norcamphor, étudié par Holderna-Natkaniec et al.[59] en utilisant la diffraction de neutrons et le scattering inélastique incohérent dans une plage de température allant de 10 K à 290 K. Dans leur étude par diffraction de neutrons, les paramètres de maille ne sont pas reportés il n'y aucune tentative de résoudre la structure, à cause de la basse résolution du diffractogramme. Nous avons pu résoudre ces structures grâce à la haute résolution de la ligne ID31 ; mais des incertitudes, dues a la présence d'une orientation préférentielle, persistent pour la phase à plus basse température. Enfin, les autres composées étaient

complètement inconnus. Lorsque nous n'avons pas pu résoudre les structures, les problèmes étaient imputable à la qualité des l'échantillons (qui présentaient une forte granularité ou une orientation préférentielle des grains). Dans ce cas, nous avons tenté de simuler la présence de l'orientation préférentielle, mais essayer de simuler en même temps le modèle et un effet qui va changer la valeur de l'intensité s'avère trop complexe, du fait de leur corrélation. Un spinner de type Gandolfi pourrait constituer une solution pour ces composés : l'échantillon devrait alors être orienté dans le plus grand nombre de directions possibles pendant l'acquisition des données. Un bilan exhaustif des résultats obtenus est présenté à la fin du chapitre.

Le chapitre 4 a été consacré à une étude sur les opïoides neuropeptidiques. La structure la plus complexe que nous avons pu résoudre était la molécule de Glycine-Glycine-Phénylalanine, ne possédant pas d'atomes lourds. Cependant, atteindre ce résultat n'a nécessité que quelques heures de calcul sur un processeur Pentium 4, 1.5 Ghz. Jusqu'ici, nous n'avons pas été confronté à des limites de la simulation dans le cas où l'échantillon était de bonne qualité. Cependant, il est impossible de prédire jusqu'à quel degré de complexité de structure la simulation peut être efficace: des expériences supplémentaires doivent être réalisées, sur des échantillons plus complexes. Il conviendrait alors de se rappeler que plus la structure à analyser sera complexe, plus l'espace N-dimensionnel qu'il faudra explorer le sera, et qu'il sera plus difficile d'être sûr de la validité du résultat obtenu (notamment parce qu'il pourrait exister plusieurs minima locaux très proches et difficilement distinguables du minimum global). Des fonctions de pénalité peuvent permettre de trouver la bonne solution, au prix d'un allongement du temps de calcul.

Le chapitre 5 traite de la localisation du benzène dans des zéolithes X au sodium. Nous y avons montré que la simulation remplace avantageusement le calcul des cartes de Fourier et la recherche manuelle des atomes manquants, surtout dans les cas où il existe du désordre (rappelons que même dans le travail de Aubert et al.[79], sur monocristal, il n'a pas été possible de localiser le benzène manuellement). Dans tous les cas nous pensons que, lors de l'utilisation de cartes de densité, il est préférable de calculer la carte d'entropie maximale (ce qui, dans les cas présentés dans ce travail, ne demandait qu'une minute de calcul avec un

processeur Pentium 4.15 GHz). Ces cartes sont dépendantes du modèle de la même manière que les cartes de Fourier car le set de F_{obs} utilisé pour calculer les cartes d'entropie maximale est le même mais dans ce dernier cas les erreurs de troncature sont évitées. Il faut également accorder une attention toute particulière à la recherche des détails les plus fins, comme explicité à la fin du paragraphe de l'entropie. Ceci constitue un thème de recherche encore ouvert [40,41]

Nous avons montré que presque tous les échantillons ont été indexés avec succès grâce à la haute résolution disponible sur la ligne ID31. Cependant, les impuretés posent des problèmes pour l'indexation car, à cause de l'importance du rayonnement incident, toute trace d'impureté sera détectée. Les programmes d'indexation peuvent prendre ceci en compte, mais il est quelquefois délicat de déterminer la bonne solution parmi celles alors proposées.

La rapide dégradation des échantillons (quelques minutes) soumis au rayonnement X a posé des problèmes expérimentaux. Il fallait, au cours d'une même expérience, translater le capillaire par rapport au faisceau X de très nombreuses fois, afin de toujours sonder une partie non altérée de l'échantillon. Si une telle manipulation était possible, bien que laborieuse, à température ambiante, elle devenait quasi impossible techniquement lorsque le flux d'azote cryogénique arrosait le capillaire.

La solution trouvée consiste en un nouvel appareillage permettant au capillaire d'être translaté automatiquement, y compris jusqu'à l'intérieur du tuyau d'arrivée d'azote. De plus, nous avons mis en place un système robotisé permettant de changer le capillaire si celui ci est endommagé. Ceci explique pourquoi il nous a été difficile, durant les deux dernières années de ce travail, d'accumuler des données à différentes températures, afin de tirer partie de l'anisotropie de l'expansion thermique. De cette façon, une approche de type pseudo-monocristalline peut être appliquée aux diffractogrammes de poudre plus difficiles à indexer. Le même principe peut être utilisé pour partitionner le recouvrement des intensités diffractées, afin d'extraire de meilleures intensités pour résoudre des structures plus complexes. Cependant, nous avons montré que, dans tous les cas, la simulation était de suffisamment bonne qualité pour résoudre toutes les structures que nous avons étudiées, en particulier grâce aux programmes FOX et TOPAS.

Pour des études futures, il serait intéressant de poursuivre ce travail sur les neuropeptides ou sur d'autres échantillons biologiques cristallisés sans utilisation de solvant autres que l'eau, afin de chercher d'éventuelles différences avec des travaux sur monocristaux ou des structures inconnues. On peut également chercher à exploiter l'anisotropie de l'expansion thermique, puisque de nombreux problèmes techniques ont été résolus sur la ligne. Enfin, une amélioration prometteuse du dispositif serait d'utiliser un détecteur bidimensionnel pour vérifier la cristallinité de l'échantillon et s'assurer de l'absence d'orientation préférentielle ou de granularité, qui affectent les intensités mesurées et compliquent les résolutions de structures. On pourrait également développer un détecteur linéaire monodimensionnel, pour pouvoir collecter des données rapidement et sur un large espace angulaire, pour les échantillons très sensibles aux radiations X, à des fins d'investigation détaillée des transitions de phases de l'échantillon.

Appendix

Final refined distance and angles values for the zeolite with the chemisorbed benzene:

1.44 B/SPC RT	Si-tetrahedra (Å)	Al-tetrahedra (Å)
T-O1	1,614(8)	1,680(9)
T-O2	1,624(9)	1,681(13)
T-O3	1,633(8)	1,689(8)
T-O4	1,644(8)	1,706(8)

2.07 B/SPC RT	Si-tetrahedra (Å)	Al-tetrahedra (Å)
T-O1	1.622(8)	1.670(9)
T-O2	1.629(8)	1.697(9)
T-O3	1.638(8)	1.693(8)
T-O4	1.643(8)	1.704(8)

2.88 B/SPC 1.5K	Si-tetrahedra (Å)	Al-tetrahedra (Å)
T-O1	1,637(8)	1,677(9)
T-O2	1,627(9)	1,707(12)
T-O3	1,6292(8)	1,683(8)
T-O4	1,638(8)	1,696(8)

1.44 B/SPC RT	Si-tetrahedra (°)	Al-tetrahedra (°)
O1-T-O2	114,3(9)	111,9(8)
O1-T-O3	110,6(8)	112,4(8)
O1-T-O4	106,1(9)	109,1(9)
O2-T-O3	104,4(8)	104.3(8)
O2-T-O4	106,7(8)	107,3(8)
O3-T-O4	114,8(8)	111,6(4)

2.07 B/SPC RT	Si-tetrahedra (°)	Al-tetrahedra (°)
O1-T-O2	118.5(5)	110.4(5)
O1-T-O3	111.4(4)	113.4(5)
O1-T-O4	106.6(4)	109.0(4)
O2-T-O3	103.4(4)	104.1(4)
O2-T-O4	105.3(4)	106.8(4)
O3-T-O4	111.5(5)	112.8(5)

2.88 B/SPC 1.5K	Si-tetrahedra (°)	Al-tetrahedra (°)
O1-T-O2	112,6(11)	113,9(11)
O1-T-O3	112,3(9)	111,4(7)
O1-T-O4	105,9(11)	108,9(11)
O2-T-O3	105,4(9)	104,1(11)
O2-T-O4	106,9(9)	106,1(9)
O3-T-O4	113,5(9)	112,1(10)

1.44 B/SPC RT	(°)
Si-O1-Al	132,5(5)
Si-O2-Al	147,4(5)
Si-O3-Al	139,7(5)
Si-O4-Al	145,5(5)

2.07 B/SPC RT	(°)
Si-O1-Al	132.4(4)
Si-O2-Al	144.9(4)
Si-O3-Al	139.3(4)
Si-O4-Al	145.9(4)

2.88 B/SPC 1.5K	(°)
Si-O1-Al	133,8(5)
Si-O2-Al	145,5(4)
Si-O3-Al	140,1(5)
Si-O4-Al	143,8(5)

1.44 B/SPC RT		2.07 B/SPC RT		2.88 B/SPC 1.5K	
Na1(0,0,0)		Na1(0,0,0)		Na1(0,0,0)	
Na1-Al	3,484(16)	Na1-Al	3.501(7)	Na1-Al	3,502(18)
Na1-O3	2.698(6)	Na1-O3	2.716(5)	Na1-O3	2,736(5)

1.44 B/SPC RT	(Å)	2.07 B/SPC RT	(Å)	2.88 B/SPC 1.5K	(Å)
Na2(0.0449(8), = ,=)		Na2(0,0477(2), = ,=)		Na2(0,0474(3), = ,=)	
Na2-Al	3.179(15)	Na2-Al	3.196(6)	Na2-Al	3,207(18)
Na2-O2	2.876(6)	Na2-O2	2.878(4)	Na2-O2	2,889(6)
Na2-O3	2.194(8)	Na2-O3	2.210(4)	Na2-O3	2,223(5)

1.44 B/SPC RT	(Å)	2.07 B/SPC RT	(Å)
Na2'(0.0527(19),=,=)		Na2'(0,0499(16),=,=)	
Na2i-Al	3.244(24)	Na2i-Al	3.213(6)
Na2i-O2	2.895(12)	Na2i-O2	2.882(4)
Na2i-O3	2.270(25)	Na2i-O3	2.230(4)

1.44 B/SPC RT	(Å)	2.07 B/SPC RT	(Å)
Na3(0.2272(5),=,=)		Na3(0.2313(4),=,=)	
Na4-Al	3.233(13)	Na4-Al	3.261(6)
Na4-O2	2.364(7)	Na4-O2	2.380(4)
Na4-O4	2.872(6)	Na4-O4	2.886(5)

1.44 B/SPC RT	(Å)	2.07 B/SPC RT	(Å)	2.88 B/SPC 1.5K	(Å)
Na3'(0.2366(11),=,=)		Na3'(0,2392(6),=,=)		Na3'(0.2369(3),=,=)	
Na4i-Al	3.326(19)	Na4i-Al	3.356(6)	Na4-Al	3,338(15)
Na4i-O2	2.491(18)	Na4i-O2	2.508(4)	Na4-O2	2,458(6)
Na4i-O4	2.911(10)	Na4i-O4	2.936(5)	Na4-O4	2,923(6)

Interaction (Å) cations site III' with benzene, 1.44 Benz/SPC

Na6_O1	3.11(9)	Na7_SI	3.24(8)
Na6_Na3	4.18(9)	Na7_O1	2.69(9)
Na6_Na3'	3.88(10)	Na7_O4	2.53(9)
Na6_Na5	3.71(18)	Na7_Na6	2.14(12)
Na6_Na7	2.14(12)	Na7_Na6	4.20(15)
Na6_Na7	2.88(14)	Na7_Na6	2.88(14)
Na6_Na7	4.20(15)	Na7_Na7	2.77(10)
Na6_C1	1.94(10)	Na7_C1	2.62(10)
Na6_D1	1.81(9)	Na7_C1	2.67(9)
Na6_C2	2.71(10)	Na7_C1	1.68(8)
Na6_C2	2.84(10)	Na7_D1	2.86(10)
Na6_D2	2.89(10)	Na7_D1	2.28(9)
Na6_D2	2.48(10)	Na7_C2	2.43(10)
Na6_C3	1.90(10)	Na7_C2	2.32(8)
Na6_C3	2.95(10)	Na7_C2	2.85(9)
Na6_D3	1.36(9)	Na7_C2	2.95(10)
Na6_C4	2.79(9)	Na7_D2	1.74(10)
Na6_C4	1.83(10)	Na7_D2	1.64(9)
Na6_D4	2.33(9)	Na7_D2	3.03(9)
Na6_D4	3.03(8)	Na7_C3	2.81(8)
Na6_D4	2.17(9)	Na7_C3	2.55(8)
Na6_C5	1.32(9)	Na7_C3	2.28(10)
Na6_D5	1.36(8)	Na7_D3	2.16(9)
Na6_C5	2.38(10)	Na7_D3	1.66(9)
Na6_C5	2.31(9)	Na7_C4	2.57(7)
Na6_D6	2.64(10)	Na7_C4	3.15(10)
Na6_D6	2.51(9)	Na7_D4	2.89(8)
Na6_D6	2.91(8)	Na7_C5	3.09(9)
Na7_Al	3.54(10)	Na7_C5	2.94(8)
Na7_Si1	3.54(10)	Na7_C5	2.42(10)
Na7_D6	2.61(9)	Na7_C5	2.44(8)

Interaction (Å) cations site III' with benzene, 2.08 Benz/SPC

Na6_Na3	4.16(8)	Na7_Na6	3.34(11)
Na6_Na3'	3.90(8)	Na7_Na6	4.12(11)
Na6_Na5	1.89(14)	Na7_Na6	3.51(10)
Na6_Na6	3.45(11)	Na7_Na7	2.75(7)
Na6_Na6	3.34(11)	Na7_C1	1.72(7)
Na6_Na6	3.51(10)	Na7_C1	2.83(8)
Na6_Na6	4.12(11)	Na7_C1	2.58(7)
Na6_C1	2.60(8)	Na7_D1	1.23(7)
Na6_C1	2.34(9)	Na7_D1	2.31(7)
Na6_D1	2.72(7)	Na7_C2	2.50(7)
Na6_D1	2.17(8)	Na7_C2	2.84(7)
Na6_D3	2.79(8)	Na7_C2	3.13(7)
Na6_C4	2.49(7)	Na7_C2	2.19(7)
Na6_C4	2.21(8)	Na7_D2	2.71(7)
Na6_D4	1.43(7)	Na7_D2	2.08(7)
Na6_D4	1.84(8)	Na7_D2	1.43(7)
Na6_C5	2.77(8)	Na7_C3	2.53(7)
Na6_C5	2.18(7)	Na7_C3	3.03(7)
Na6_D5	3.01(7)	Na7_C3	2.85(7)
Na6_D5	1.79(7)	Na7_D3	1.55(7)
Na6_C6	1.76(8)	Na7_D3	2.52(7)
Na6_C6	2.48(8)	Na7_D3	2.86(7)
Na6_D6	2.41(8)	Na7_C4	2.93(7)
Na7_Al	3.20(7)	Na7_C4	3.16(7)
Na7_Si1	3.20(7)	Na7_D4	2.98(7)
Na7_SI	3.47(7)	Na7_C5	2.86(7)
Na7_O1	2.77(7)	Na7_C6	2.65(7)
Na7_O4	2.46(7)	Na7_C6	2.69(8)
Na7_Na6	3.45(11)	Na7_D6	2.96(7)

Interaction (Å) cations site III' with benzene, 2.88 Benz/SPC

Na6_Na6	1.44(16)	Na7_Na6	3.65(14)
Na6_Na7	3.65(17)	Na7_Na6	2.94(14)
Na6_Na7	3.65(14)	Na7_Na6	3.91(17)
Na6_Na7	3.91(17)	Na7_Na7	2.65(11)
Na6_Na7	2.94(14)	Na7_Na7	2.65(11)
Na6_C1a	2.40(10)	Na7_C1a	2.46(10)
Na6_C1a	3.20(8)	Na7_C2a	3.18(9)
Na6_D1a	1.89(9)	Na7_C3a	2.78(10)
Na6_D1a	2.39(8)	Na7_D3a	2.98(11)
Na6_C2a	2.16(9)	Na7_C4a	2.66(9)
Na6_C2a	2.97(7)	Na7_C4a	2.85(11)
Na6_D2a	2.47(9)	Na7_C4a	1.39(10)
Na6_C3a	1.33(11)	Na7_D4a	2.96(10)
Na6_C3a	2.27(8)	Na7_D4a	1.27(10)
Na6_D3a	2.14(8)	Na7_C5a	1.75(9)
Na6_C4a	2.35(12)	Na7_C5a	2.68(10)
Na6_C4a	2.65(10)	Na7_C5a	1.39(9)
Na6_D4a	2.74(11)	Na7_D5a	2.10(10)
Na6_D4a	2.92(10)	Na7_D5a	2.68(9)
Na6_C6a	2.91(12)	Na7_D5a	1.27(9)
Na6_D6a	2.95(12)	Na7_C6a	2.83(10)
Na7_SI	3.36(9)	Na7_C6a	1.64(9)
Na7_O1	2.58(9)	Na7_C6a	2.78(9)
Na7_O4	2.88(10)	Na7_D6a	1.96(10)
Na7_Na6	3.65(17)	Na7_D6a	2.06(10)

Bibliography

1. Giacovazzo, C., *Fundamentals of crystallography 2nd edition*. IUCr Text in Crystallography: Oxford University Press (2002).
2. Steward, E.G., *Fourier optics an introduction 2nd edition*. Hellis Horwood series in physics (1987).
3. Taylor and Lipson, *Optical transform*. Bell, London (1964).
4. Rehg, J., *Filtering*, in *www.cc.gatech.edu/classes/AY2004/cs4495_fall/Materials/22-Filtering2_CS4495_10-10-03.ppt*. (2003).
5. Jenkins, R. and R.L. Snyder, *Introduction to X-ray powder diffractometry*. Jhon Wiley & Sons, Inc. 1996.
6. David, W.I.F., et al., *Structure determination from powder diffraction data*. IUCr Text in Crystallography: Oxford University Press (2002).
7. Rietveld, H.M., *Line profiles of neutron powder-diffraction peaks for structure refinement*. Acta Cryst, 1967. **22**: p. 151-152.
8. Rietveld, H.M., *A profile refinement method for nuclear and magnetic structures*. J. Appl. Crystallogr, 1969. **2**: p. 65-57.
9. Young, R.A., *The Rietveld method*. IUCr Text in Crystallography: Oxford University Press (1995).
10. Wolff, P.M., *The definition of the indexing figure of merit M20*. Journal of Applied Crystallography, 1972. **5**(3): p. 243.
11. Shirley, R., *The Crysfire 2002 System for Automatic Powder Indexing: User's Manual*. 2002.
12. Boultif, A. and D. Louer, *Powder pattern indexing with the dichotomy method*. Journal of Applied Crystallography, 2004. **37**(5): p. 724-731.
13. Visser, J., *A fully automatic program for finding the unit cell from powder data*. Journal of Applied Crystallography, 1969. **2**(3): p. 89-95.
14. Werner, P.-E., L. Eriksson, and M. Westdahl, *TREOR, a semi-exhaustive trial-and-error powder indexing program for all symmetries*. Journal of Applied Crystallography, 1985. **18**(5): p. 367-370.
15. Taupin, D., *A powder-diagram automatic-indexing routine*. Journal of Applied Crystallography, 1973. **6**(5): p. 380-385.
16. Kohlbeck, F. and E.M. Horl, *Indexing program for powder patterns especially suitable for triclinic, monoclinic and orthorhombic lattices*. Journal of Applied Crystallography, 1976. **9**(1): p. 28-33.
17. Le Bail, A., *McMaille program*, in *http://www.cristal.org/McMaille*. 2002.

18. Coelho, A., *Indexing of powder diffraction patterns by iterative use of singular value decomposition.* Journal of Applied Crystallography, 2003. **36**(1): p. 86-95.

19. Le Bail, A., *The Rietveld method using an experimental profile convoluted by adjustable analytical function.* Acta Crystallographica Section A, 1984. **40**(s1): p. C369.

20. Pawley, G., *Unit-cell refinement from powder diffraction scans.* Journal of Applied Crystallography, 1981. **14**(6): p. 357-361.

21. Waser, J., *Least-squares refinement with subsidiary conditions.* Acta Crystallographica, 1963. **16**(11): p. 1091-1094.

22. Kirkpatrick, S., C.D. Gelatt, and M.P. Vecchi, *Optimization by simulated annealing.* Science, 1983. **220**(4598): p. 671-680.

23. Metropolis, N., et al., *Equation of state calculation by fast computing machines.* The journal of chemical physics, 1952. **21**(6): p. 1087-1092.

24. Shankland, K., et al., *Molecular, crystallographic and algorithmic factors in structure determination from powder diffraction data by simulated annealing.* Journal of Applied Crystallography, 2002. **35**(4): p. 443-454.

25. Falcioni, M. and M.W. Deem, *A biased Monte Carlo scheme for zeolite structure solution.* The Journal of Chemical Physics, 1999. **110**(3): p. 1754-1766.

26. Favre-Nicolin, V. and R. Cerny, *FOX, `free objects for crystallography': a modular approach to ab initio structure determination from powder diffraction.* Journal of Applied Crystallography, 2002.

27. Polak, E., *§2.3 Computational Methods in Optimization.* New York: Academic Press. 1971.

28. Barrett, R., et al., *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd ed.* Philadelphia, PA: SIAM,. 1994.

29. Golub, G. and D. O'Leary, *Some history of the conjugate gradient and Lanczos methods.* SIAM Rev., 31. 1989. 50-102.

30. Cambridge, *http://www.library.cornell.edu/nr/bookcpdf/c2-3.pdf.* 1992.

31. Coelho, A., *Whole-profile structure solution from powder diffraction data using simulated annealing.* Journal of Applied Crystallography, 2000. **33**(3 Part 2): p. 899-908.

32. Shewchuk, J.R., *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain.* http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.ps. 1994.

33. Jaynes, E.T., *Information theory and statistical mechanics.* Physical Review, 1957. **106**: p. 620.

34. Jaynes, E.T., *Information theory and statistical mechanics II.* Physical Review, 1957. **108**: p. 171.

35. Gull, S.F. and G.J. Daniell, *Image reconstruction from incomplete and noisy data.* Nature, 1978. **272**: p. 686.

36. M. Takata, E. Nishibori, and M. Sakata, *Charge density studies utilizing powder diffraction and MEM. Exploring of high Tc superconductors, C60 superconductors and manganites.* Zeitschrift für Kristallographie, 2001. **216**(2): p. 71.

37. Palatinus, L., *Maximum Entropy Method in Superspace Crystallography*, in *http://opus.ub.uni-bayreuth.de/volltexte/2004/80/index.html.* 2003.

38. Sivia, D.S., *Data Analysis a Bayesian Tutorial.* Oxford University Press (1996).

39. F. Izumi and R. A. Dilanian, *Recent Research Developments in Physics," Vol. 3, Part II*. Transworld Research Network, Trivandrum (2002), pp. 699-726.

40. Kumazawa, S., M. Takata, and M. Sakata, *On the single-pixel approximation in maximum-entropy analysis*. Acta Crystallographica Section A, 1995. **51**(1): p. 47-53.

41. S. van Smaalen, L. Palatinus, and M. Schneider, *The maximum-entropy method in superspace*. Acta Crystallographica Section A, 2003. **59**(5): p. 459-469.

42. L. Palatinus and S. van Smaalen, *The generalized F constraint in the maximum-entropy method - a study on simulated data*. Acta Crystallographica Section A, 2002. **58**(6): p. 559-567.

43. Margaritonodo, G., *Introduction to synchrotron radiation*. Oxford University Press (1988).

44. Mason, R., *ESRF beamline handbook*. European Synchroton Radiation Facility (1994).

45. J.L. Hodeau, et al., *Nine crystal multi-analyzer stage for high resolutiuon powder diffraction between 6 and 40 keV*. Part of the SPIE confernce on crystal and multilayer optics, 1998. **3448**: p. 353.

46. Kitaigorodsky, A.I., *Molecular crystals and molecules*. Academic Press: Physical chemistry.

47. T. Steiner, *The hydrogen bonod in the solid state*. Angewandte Chemie (Int. Eng. Ed.), 2002. **41**: p. 48-76.

48. Timmermans, J., *Plastic crystals: A historical review*. Journal of Physics and Chemistry of Solids, 1961. **18**: p. 1.

49. D. Fox, M. Labes, and A. Weissberger, *Physics and chemistry of the organic solid state*. Interscience publisher (1963).

50. C. C. Mjojo, *Order-Disorder Phenomena. Order-Disorder Phase Equilibria in D- and L-systems of Camphor and Related Compounds*. Journal of the Chemical Society. Faraday Transactions, 1979. **75**: p. 692-703.

51. M. Brunelli, A. N. Fitch, and A. Mora, *Low-Temperature Crystal Structure of S-camphor Solved from Powder Synchrotron X-ray Diffraction Data by Simulated Annealing*. Journal of Solid State Chemistry, 2002. **163**: p. 253-258.

52. R. M. Paroli, D.F. Gilson, and L. S. Butler, *Specroscopic adn differential scanning calorimetric studies of the phase transition in oxanorborane*. Canadian Journal of Chemistry, 1998. **76**: p. 1365-1370.

53. A. C. Larson and R. B. von Dreele, *GSAS, General Structure Analysis System.*, Los Alamos Natl. Lab. Rep. LA-UR-86-784 (1987).

54. Thompson, P., D.E. Cox, and J.B. Hastings, *Rietveld refinement of Debye-Scherrer synchrotron X-ray data from Al2O3*. Journal of Applied Crystallography, 1987. **20**(2): p. 79-83.

55. Finger, L.W., D.E. Cox, and A.P. Jephcoat, *A correction for powder diffraction peak asymmetry due to axial divergence*. Journal of Applied Crystallography, 1994. **27**(6): p. 892-900.

56. R.A. Creswell, *Microwave spectrum, strucure, and dipole moment of 7-Oxabicyclo[2.2.1]heptane*. Journal of Molecular Spectroscopy, 1975. **56**: p. 133-137.

57. Jarvinen, M., *Application of symmetrized harmonics expansion to correction of the preferred orientation effect*. Journal of Applied Crystallography, 1993. **26**(4): p. 525-531.

58. Hall, M.M., et al., *The approximation of symmetric X-ray peaks by Pearson type VII distributions*. Journal of Applied Crystallography, 1977. **10**(1): p. 66-68.

59. Holderna-Natkaniec, K., I. Natkaniec, and V. Khavryutchenko, *Neutron-scattering study of phase transition in norbornylene and -norcamphor.* Physica B: Condensed Matter, 1997. **241-243**: p. 478-480.

60. P.S. Katz, *Beyond Neurotransmission: Neuromodulation and Its Importance for Information Processing.* Oxford University Press (1999).

61. A. Stephenson and A.J. Turner, *Amino Acid Neurotransmission.* Portland Press 1997.

62. Shwartz , J.H., *neurotransmitters.* encyclopedia of life sciences Nature publishing group, 2001.

63. J. Hughes, et al., *Identification of two related pentapeptide from the brain with potent opiate agonist activity.* Nature, 1975. **258**: p. 577.

64. Karle, I.L., et al., *[Leu5]enkephalin: four cocrystallizing conformers with extended backbones that form an antiparallel b-sheet.* Acta Crystallographica Section B, 1983. **39**: p. 625-637.

65. J.F. Griffin, et al., *The crystal structures of [Met]enkephalin and a third form of [Leu]enkephalin: Obseravtion of a novel pleated b-sheet.* Proceedings of the National Academy of Science of the U.S.A., 1986. **83**: p. 3272-3276.

66. Smith, G.D. and J.F. Griffin, *Conformation of [Leu]enkephalin from X-ray diffraction: features important for recognition at opiate receptor.* Science, 1978. **199**: p. 1214-1216.

67. A. Aubry, et al., *A crystal molecualr conformation of lucine-enkephalin related to the morfine moelcule.* Biopolymers, 1989. **28**: p. 27-40.

68. Prange, T. and C. Pascard, *Opiaces peptidiques: structure et conformation de deux fragments d'enkephalines, Tyr-Gly-Gly-Phe et G1y-Gly-Phe-Leu.* Acta Crystallographica Section B, 1979. **35**(8): p. 1812-1819.

69. Carson, W.M. and M.L. Hackert, *Crystal structure of l-tyrosyl-glycyl-glycine monohydrate, the N-terminal tripeptide of the enkephalins.* Acta Crystallographica Section B, 1978. **34**(4): p. 1275-1280.

70. Pichon-Pesme, V., et al., *Electron density and electrostatic properties of two peptide molecules: tyrosyl-glycyl-glycine monohydrate and glycyl-aspartic acid dihydrate.* Acta Crystallographica Section B, 2000. **56**(4): p. 728-737.

71. Engh, R.A. and R. Huber, *Accurate bond and angle parameters for X-ray protein structure refinement.* Acta Crystallographica Section A, 1991. **47**(4): p. 392-400.

72. A. Dyer, *An introduction to zeolite molecular sieves.* John Wiley & Sons (1988).

73. Mortier, W.J., *In Compilation of Extra-Framework Sites in Zeolites.* Butterworth Scientific Ltd.: Guildford, U.K., 1982.

74. Bennett, J.M. and J.V. Smith, *Positions of cations and molecules in zeolites with the faujasite-type framework I. Dehydrated Ca-exchanged faujasite.* Materials Research Bulletin, 1968. **3**(8): p. 633-642.

75. Olson, D.H., *A Reinvestigation of the Crystal Structure of the Zeolite Hydrated NaX.* Journal of Physical Chemistry, 1970. **74**: p. 2758.

76. Olson, D.H., *The crystal structure of dehydrated NaX.* Zolites, 1995. **15**: p. 439-443.

77. Bull, L.M., et al., *A neutron diffraction study of acid sites in H-SAPO-37.* Journal of the Chemical Society. Chemical Communications, 1993. **15**: p. 1196.

78. Hriljac, J.A., et al., *Powder Neutron Diffraction and 29Si MAS NMR Studies of Siliceous Zeolite-Y*. Journal of Solid State Chemistry, 1993. **106**(1): p. 66-72.

79. A. N. Fitch, H. Jobic, and A. Renouprez, *Localisation of benzene in sodium-Y zeolite by powder neutron diffraction*. Journal of Physical Chemistry, 1986. **90**: p. 1311-1318.

80. Aubert, E., et al., *Adsorption of aromatic molecules in NaX zeolite: trying a structure determination by single-crystal X-ray diffraction*. Acta Crystallographica Section A, 2000. **56**(s1): p. s391.