Tutor

Dr. Pere Alemany Cahner
Departament de Ciència de Materials i Química Física.

Grau de Química

Treball Final de Grau

Crystal Packing of Long-chain Aliphatic Molecules: A Continuous Symmetry Measures Approach.

Empaquetament cristal·lí de compostos amb cadenes alifàtiques llargues: una aproximació mitjançant mesures contínues de simetria.

Laura Sánchez Muñoz

January 2021

UNIVERSITAT DE BARCELONA

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

L'aspecte més trist de la vida actual és que la ciència guanya en coneixement més ràpidament que la societat en saviesa.

Isaac Asimov

Voldria agrair tot l'esforç realitzat pel Dr. Pere Alemany que m'ha guiat i ajudat a poder realitzar satisfactòriament aquest treball malgrat les circumstàncies actuals.

REPORT

CONTENTS

1. SUMMARY

Bulk crystals of normal Alkanes n-C_nH_{2n+2} and Langmuir monolayers of alkane derivates show similar phase diagrams, including rotator phases and the low-temperature herringbone structure. By varying the temperature, the interactions between neighboring chains will be altered due to the increase in atomic movement, thus modifying the structure in which the chains pack. In this project, a study will be carried out to be able to compare the theoretical and experimental packing modes for normal alkanes. To carry out these comparisons we made a search in a crystallographic database (Cambridge Structural Database) for the structures needed to perform a crystallographic analysis as well as some molecular mechanics calculations for ideal infinite aliphatic chains. Once the cell parameters and atomic positions have been obtained for different compounds, a geometric analysis has been made to obtain the packing modes of the alkanes and describe their changes during phase transitions.

Keywords: Crystal Packing, CSD database, Crystallography, Alkanes, Molecular Structure.

2. RESUM

Els cristalls d'alcans normals n-C_nH_{2n+2} i les monocapes de Langmuir de derivats d'alcans mostren diagrames de fase semblants, incloent les diverses fases rotadores i l'estructura d'espiga de baixa temperatura. En variar la temperatura les interaccions entre cadenes veïnes es veuran alterades degut a l'augment del moviment atòmic i això provocarà canvis estructurals a l'hora d'empaquetar-se. En aquest treball s'ha realitzat un estudi comparant els empaquetaments d'alcans normals, tant teòrics com experimentals. Per poder dur a terme aquestes comparacions s'ha fet prèviament una recerca en una base de dades cristal·logràfica (Cambridge Structural Database) i càlculs de mecànica molecular per a cadenes ideals infinites per obtenir les estructures necessàries per a procedir a realitzar un anàlisi cristal·logràfic. Un cop obtinguts els paràmetres de cel·la i les posicions atòmiques s'ha realitzat una anàlisi geomètrica per obtenir els diferents modes d'empaquetament dels alcans i descriure les transicions de fase que hi tenen lloc.

Paraules clau: Empaquetament cristal·lí, Base de dades CSD, Cristal·lografia, Estructura molecular, Alcans.

3. INTRODUCTION

Paraffin wax (or petroleum wax) is a soft colorless solid derived from petroleum that consists of a mixture of hydrocarbon molecules containing between twenty and forty carbon atoms. In chemistry, the term paraffin is often used as a synonym for normal alkanes, that is, aliphatic hydrocarbons with the general formula C_nH_{2n+2}. The name paraffin comes from the latin *Parum affinis* which means low affinity, because alkanes aren't very reactive compounds. Normal alkanes have a simple structure consisting in a chain formed by (*n*-2) CH_2 groups joined by simple C-C bonds which is capped by two terminal CH_3 groups. In the last 50 years, paraffin crystals have been the object of intense study because a large number of organic and biological compounds contain fragments of aliphatic chains and paraffins are simple models to understand the behavior of other more complex systems, such as liquid crystal phases, membranes or compounds formed by different layers. It has been found that some physical properties of paraffins such as the density of the liquid, the boiling point, or the melting point depend on the molecular weight or what is equivalent, on the number of carbon atoms in the compound, with some interesting differences depending on if this number is even or odd. Paraffins have low dielectric constant values, are apolar compounds, with extremely low electrical conductivity and solubility in water. In the solid state, paraffins crystalize in a diverse number of polymorphs whose structures are mainly dictated by weak non-covalent van der Waals interactions between individual molecules arranged in layers.

In this work we will focus our attention on the formation of crystals of long-chain n-alkanes containing more than six carbon atoms, and particularly on the way that long aliphatic chains pack in the solid state. In this respect, a good model for understanding the effects of lateral interactions between neighboring chains in the solid state is polyethylene, $-(CH_2)_n-$, which is a good model for a regular infinite aliphatic chain. Before we start to explain in more detail the different packing modes and interactions for infinite and finite aliphatic chains, we will introduce some basic concepts about packing in molecular crystals.

3.1. MOLECULAR PACKING CRYSTALS

A molecular crystal can be defined as a crystal formed by the tight packing of individual molecules, which may be identified as a group of atoms in a crystal in which the distance to each atom in the group is significantly less than the distances to atoms in other adjacent groups. In general, there is not a unique way of packing equal objects (molecules) to form a regular crystalline arrangement and alkane crystals are often found in different alternative structures, also known as *polymorphs*, that exhibit a similar stability but differ in the molecular packing arrangement, restrictions on twisting of the chains, or on the rotating movement of the chains around their axes.

The general geometric principles of packing of molecules to form crystals has been thoroughly described by Kitaigorodskii in references [1,2] from which we will use several concepts later. By comparing numerous crystal structures, Kitaigorodskii formulated the *principle of close packing*, according to which in the absence of strongly directional intermolecular interactions such as hydrogen bonds, molecules will tend to form a packing with the highest possible density, that is, with the largest fraction of space occupied by atoms in a cell (*packing factor* or *packing coefficient*). The higher the packing coefficient, the more stable that structure will be and will thus have a higher tendency to form. The orientation of the molecules in a crystal is therefore conditioned by the short distances between the closest atoms of neighboring molecules which should be as short as possible, but avoiding the strong repulsive forces that arise when two atoms are forced to be too close together. In the case of organic molecules, the different packing alternatives are often defined by the interactions between the hydrogen atoms on the external surface of the molecules, and a general guiding principle to predict the optimal packing is to try to fit two molecules together avoiding intermolecular H···H contacts with distances below twice the van der Waals radius of H, which is normally taken to be around 1.2Å.

Next, we will introduce some specific features of the packing of long-chain molecules to later develop these ideas in more detail considering first the ideal case of infinite aliphatic chains (polyethylene) and then that of finite chains (paraffins).

3.2. PACKING OF LONG-CHAIN MOLECULES

In the absence of directional interactions such as hydrogen bonds, the principal factor determining the optimal packing is molecular shape. In crystals, long-chain aliphatic molecules adopt a linear extended configuration for which packing in layers of chains with parallel long axes is favored and

the possible crystalline structures that they may form can be obtained from a simple geometric analysis of their structure. For simple cylindrical objects with a perfect circular section the most compact packing mode leads to a hexagonal lattice with all objects having their long axes aligned in the perpendicular direction (Figure 1 a). An alternative arrangement is based on a square net (Figure 1 b).

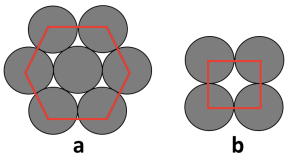

Figure 1. Compact packing modes for cylindrical objects: a) hexagonal packing and b) square packing.

The packing factors for these two structures can be easily calculated from simple geometric considerations. In the square packing, we have a coordination number 4, that is, each cylindrical object would be in contact with other four neighbors and the packing factor in 2D is

$$PE = \frac{area\ of\ atoms\ in\ unit\ cell}{area\ of\ unit\ cell} \qquad 1.$$

Knowing that the area of circles in the unit cell is πr^2 and the area of the unit cell is $4r^2$, a packing factor of 0.78 is obtained, that means that 78% of unit cell will be occupied while the 22% will be empty space. This value applies also to the volume occupied by cylinders of height *h* tacking into account that both the volume of each cylinder and of the unit cell is just its section multiplied by *h*, so that the height cancels out in the division to obtain the packing coefficient.

In the case of the hexagonal compact packing the coordination number is six and the area of the circles in the unit cell is $3\pi r^2$, while the area of the unit cell is $6\sqrt{3r^2}$ and we obtain a packing factor of 0.91. Therefore, comparing both factors, the hexagonal packing is more efficient than the square one since there is less empty space left when packing the objects together.

Since aliphatic chains don't have a perfect circular shape, this will give us packings which can be seen as distortions from the ideal hexagonal or square ones. An interesting question that arises is that when temperature is increased, paraffin chains are able to rotate around their long axis, giving the so-called *rotator phases* where, on average, the section of the chains becomes progressively circular, and the observed packing becomes progressively hexagonal.

Since the structure of paraffin crystals is based on layers of chains whose axes are parallel, the analysis of packing modes can be divided in two parts: the packing of the molecules single layers and the stacking of those layers to form a crystal. In this work, we were mainly interested in the influence of lateral interactions between chains and the packing modes that lead to different layer geometries. For this purpose we have used two different approaches, the search for experimental

structures in databases and the optimization of the structures using force fields to obtain crystal structures and extract the principal geometrical parameters characterizing their packing modes.

4. OBJECTIVES

The main objective of this work is to analyze the principal geometric features of crystal packing for long-chain alkanes and the changes induced in this packing by the atomic motion due to temperature. In order to carry out this objective, a search will be carried out for the structure of long-chain alkanes in the Cambridge Structural Database to obtain structural information and store it in files with CIF format. Structures for ideal infinite aliphatic chains have been obtained from molecular mechanics optimizations using the GULP program. Next, by studying the crystal structures, the packing mode of the chains in each structure is analyzed in terms of continuous symmetry and shape measures. Finally, the changes in the crystallographic parameters induced by the variation of the temperature have been studied by the changes in the continuous symmetry measures.

5.COMPUTATIONAL METHODS

As explained above, in this work we have obtained crystal structures for long-chain n-alkanes from two different sources, the Cambridge Structural Database (CSD) [3] and using the GULP program [4] to optimize the crystal structure using molecular mechanics with a force field that has been widely used to predict the structure of organic molecules and polymers. Once we have a crystal structure, consisting of a set of cell parameters and atomic positions, we have performed a geometric analysis of the packing of chains into layers by determining the points where the long axes for a given molecule and its nearest neighbors cut a plane containing equivalent atoms of each chain. The symmetry of the lateral packing of chains is based on the computation of continuous symmetry and shape measures for the arrangement of these points to see how they

differ from a perfect hexagonal symmetry. All geometric manipulations and visualization of crystal structures have been performed using the Crystal Maker software [5] and some small Python programs written on purpose for this work.

Before entering into the detailed discussion of the packing modes for paraffines we will briefly overview the principal aspects of the force field used in our calculations and the concepts of continuous shape and symmetry measures.

5. 1. THE DREIDING FORCEFIELD FOR ALIPHATIC HYDROCARBONS

For years, different force fields have been developed to be able to correctly predict the structures and dynamics of molecules. A force field is a computational method used to estimate the forces between atoms, either within molecules or between molecules. Force fields use different functional forms and parameter sets to calculate the potential energy of a system of atoms in molecular mechanics. Molecular mechanics can be used to study systems of molecules that vary in size and complexity, from biological systems or collections of materials of millions of atoms. The exact force field depends on the particular simulation program that is used, but generally, the total energy is divided into different components arising from two-body (atom-atom), three-body, and four-body interactions. The total energy of a molecule is calculated as a function of bond distances, angles and dihedral angles and the different parameters defining the force field are usually obtained by fitting procedures to balanced values of bond lengths and angles obtained from experiments or theoretical calculations of the electronic structure carried out by a software that uses ab-initio calculations such as Gaussian [6].

In this work the optimal structure for hydrocarbon crystals has been calculated by minimizing the forces using a force field that considers energy terms between 2, 3, and 4 atoms using the GULP program. We start with an initial guess geometry, which should be close to the optimal structure, and the program performs different optimization cycles in which the crystal geometry (cell parameters and atomic positions) is changed until convergence is achieved when the total stresses on the cell and the forces on the atoms are smaller than a given threshold.

Within force fields, a considerable variety of functional forms and parameter sets can be found. For example, popular choices are the MM2/MMP2 force field of Allinger et al. [7] for organic and inorganic compounds or the AMBER force field of Kollman et al. [8] for proteins and nucleic acids. In our case, we use the Dreiding force field [9] which is very useful since it has been developed

so that it can be used generically for many compounds, that is, its parameters are restricted to very simple rules. Today, the Dreiding force field is used to predict the structures and dynamics for organic, biological and inorganic main group molecules. Our choice is based on its popularity in polymer chemistry.

In order to carry out the calculations, the first step is to choose the Dreiding atom type for the different atoms present in our system. The nomenclature for atom types in Dreiding uses a five-character mnemonic tag. The first two correspond to the chemical symbol (for example H is hydrogen) and are accompanied by an underscore. The third character corresponds to hybridization or geometry; 1 = sp1, 2 = sp2, 3 = sp3. The fourth character is used to indicate the number of implicit hydrogens in united atom models. In our case, we used the H_ and the C_3 atom types for H and C atoms in our molecules to define the corresponding parameters appearing in the different energy terms.

We have used atom-atom potentials for interactions between bonded atoms within a molecule, general interatomic potentials for atom-atom interactions between non-bonded atoms either in the same molecule or on different molecules, three body potentials for the angles between bonded atoms and four-body potentials for the dihedral angles formed by sets of four bonded atoms.

Atom-atom potentials for interactions between bonded atoms are calculated from a bond stretching interaction as if it were a simple harmonic oscillator, that is, using the general expression given in [Eq. 1] with k_2 = 30.355 eVÅ^{-2} , k_3 = 0, and k_4 = 0 for any H-H, C-H or C-C bond present in the structure with the r_0 values of 0.650Å, 1.090Å and 1.530Å for H-H, C-H and C-C bonds, respectively.

$$E = \frac{1}{2}k_2(r - r_0)^2 + \frac{1}{6}k_3(r - r_0)^3 + \frac{1}{24}k_4(r - r_0)^4 \qquad 2.$$

For the case of nonbonded interactions we have used a general interatomic potential that describes the Van der Waals interactions using a Lennard-Jones potential [Eq. 2] where ε and σ are 0.659 x 10^{-3} eV and 3.195Å for H···H interactions; 0.165 x 10^{-2} eV and 3.547Å for the C···C interactions, and 0.412 x 10^{-2} eV and 3.898Å for the C···C interactions. In all three cases the values of *m* and *n* are 12 and 6, respectively.

$$E_{vdw}^{LJ} = \varepsilon\left[\left(\frac{n}{m-n}\right)\left(\frac{\sigma}{r}\right)^{m} - \left(\frac{m}{m-n}\right)\left(\frac{\sigma}{r}\right)^{n}\right] \qquad 3.$$

The three-body term describing the bending of the angle formed by three atoms contains a simple harmonic term [Eq. 4] with K_{IJK} = 4.3364432 eV rad^{-2} and an equilibrium angle θ_{J° of 109.47º for any CCC, CCH or HCH bond angle in the system.

$$E_{IJK} = \frac{1}{2}K_{IJK}\left[\theta_{IJK} - \theta_J^0\right]^2 \quad 4.$$

The torsion potential for two bonds connected through a common C-C bond forming a dihedral angle φ is given in [Eq.5] where $1/2V_{JK}$ = 0.0433 eV, n_{JK} has a value of 3, and φ_0 a value of 180/*n*.

$$E_{IJKL} = \frac{1}{2}V_{JK}\{1 - cos[n_{JK}(\varphi - \varphi_{JK}^0)]\} \quad [5.$$

5. 2. CONTINUOUS SYMMETRY MEASURES

In order to study the different distortions from the ideal hexagonal symmetry in the packings of long-chain aliphatic molecules we have used *Cosymlib* [10], a Python library developed by the Electronic Structure and Symmetry group at the Institut de Química Teòrica i Computacional de la Universitat de Barcelona (IQTCUB). *Cosymlib* can be used to compute continuous shape and symmetry measures. In our work we have used shape measures, a geometric tool that can be applied to any geometric object that is defined by a set of vertices or mass distribution function. Continuous symmetry or shape measures are based on the idea of being able to compare how much one object resembles another object with a given ideal symmetry or shape, such as, for instance a square (Figure 2).

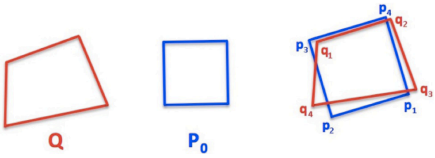

Figure 2. Superposition of the problem structure Q and the ideal square structure P_0 to obtain a continuous square measure for Q. (Image taken from Carreras, A. et al. P. Effects of Temperature on the Shape and Symmetry of Molecules and Solids. Chemistry 25, 673–691 (2019).)

If we look at figure 2, we can see a quadrangle Q (the structure to be studied) and a square P_0 (the ideal structure). To obtain the continuous shape measure of Q with respect to the square, we change P_0 using translations, rotations and an isotropic size changes until obtaining P, the square that maximizes the superposition with the problem structure Q and determine $S_P(Q)$, the continuous measure of the shape of Q respect to the ideal shape P (the square in the example). The mathematical expression that defines the shape measure is:

$$S_p(Q) = min\frac{\sum_{i=1}^{N}|q_i - p_i|^2}{\sum_{i=1}^{N}|q_i - q_0|^2} * 100 \quad 6.$$

where N refers to the number of vertices in the structures which are being compared, q_i and p_i are the position vectors of Q and P respectively and q_0 is the geometric center of the structure which we want to study. The minimization in 6. refers to the relative position, orientation and scale which must be applied to the ideal structure for a correct superposition, as well as overall permutations on vertex pairings between the two structures.

If a value S_p (Q) = 0 is obtained from equation 6 it means that the structure, we are studying has exactly the same shape as the ideal one. On the other hand, if a high value of S_p (Q) is obtained, it will have a lower resemblance to the ideal shape. The maximum value of S_p (Q) is 100 which corresponds to the non-physical situation in which all the vertices of Q collapse at the same point. When we use this type of measures for the structures of chemical compounds, they never exceed the value of S_p (Q) = 50 and, therefore, symmetry measures of 0.1 or higher are already considered to indicate chemically significant distortions.

In the case of the square used above as an example, the shape and the symmetry measure coincide because the square is the only shape with four vertices and a square D_4h symmetry. This is however only the case for a very few shapes such as the regular polygons or the five platonic solids. For a more general shape such as a rectangular prism with a regular hexagonal base, shape is however a more restrictive concept than symmetry since any hexagonal rectangular prism will have the D_{6h} symmetry irrespective of the length ratio between the hexagonal and rectangular edges. For this type of objects, a standard reference shape is included in *Cosymlib*, in this case, a rectangular hexagonal prism where all edges have the same length. Although it is possible to use this ideal shape as a reference, if we want to compare structures with significant departures from this standard reference shape, it is better to use a custom reference shape that will lead to smaller continuous shape measures. In our case we want to analyze the shape of the prism formed by the 12 points where the central axes of the six chains that surround a given chain cross the planes containing the beginning and the ending groups of the $(CH_2)_{n-2}$ chain.

Since in our work we are dealing with carbon chains from 6 to 60 carbon atoms, the longer the chain, the larger the shape measure with respect to the reference prism included in *Cosymlib*. Since using this reference shape with all edges with the same length would consider as a distortion the fact of comparing it to a prism whose height increases with the increase of carbon atoms, we have decided to use a different reference shape considering a reference hexagonal prism with the average interchain distance in experimental structure of polyethylene for the hexagonal faces and the height corresponding to the c parameter of the cell containing only a CH_2-CH_2 fragment per chain. In order to compare the prisms obtained for chains of different length with this reference shape, the prisms obtained from the experimental structures for finite alkane chains were normalized to a height of a single CH_2-CH_2 fragment, so that the “distortion” of the prism due to changes in the chain length will not affect our results.

6. PACKING OF INFINITE ALIPHATIC $(CH_2)_\infty$ CHAINS

We will start our study considering the packing of ideally infinite aliphatic chains, neglecting the effect that terminal groups may have on the arrangement of the chains in the final crystal structure. The closest experimental realization of those infinite chains is polyethylene and we will discuss its crystal structure at the end of the section. Before this, we will center our attention on the structure adopted by individual $(CH_2)_\infty$ chains when packed together. Then we will study how these chains interact with neighboring chains, and later analyze the formation of planes through the assembly of parallel chains. Finally, we will study the different possible stacking modes of different planes to form crystals.

6. 1. STRUCTURE OF A SINGLE $(CH_2)_\infty$ CHAINS

In Figure 4A we can see the most common structure adopted by long aliphatic chains in crystals. In this all-trans configuration the carbon atoms are kept on the same plane forming a zigzag chain while, on the other hand, the hydrogen atoms alternate from one side to the other of the CCC backbone with each CH_2 group lying in a plane perpendicular to the chain direction. This zigzag configuration of the chain helps to avoid repulsive interactions between the hydrogens on the same chain. Since the hydrogen atoms alternate from one side the other of the chain, the height along the chain direction corresponding to the hydrogen atoms depends on which side of the chain they are. To be able to visually differentiate hydrogen atoms at different heights we will represent them using different colors, red and blue. In this way, when it comes to visualizing the packing with more than one chain, we will have a clearer vision of which hydrogen atoms lie at the same height. The repeating unit needed to generate the whole chain by translation along the chain axis is thus formed by two CH_2 units, each one with its hydrogen atoms at opposite sides of the chain.

When discussing the different ways of packing chains together it will be more convenient to represent the structure as seen in the direction of the chains. In figure 3B the aliphatic chain is seen from above and hydrogens at either side are clearly seen to be at different heights. In this view, the carbon backbone forming the zigzag chain is seen as a single line and the orientation of the chain in a crystal will be given by the setting angle, defined as the angle formed by this C-C line and the vector corresponding to the a direction of the crystal's unit cell.

In order to build models for the different packing modes we need the approximate dimensions of each CH_2 group, which are obtained using the trigonometric relationships between the different

atomic positions based on the C-C, C-H, CCC and HCH distances and angles. For this purpose, we have used values of 1.55Å, 1.09Å, 113º and 108.1º respectively, which we have obtained from an optimization of the structure of crystalline polyethylene using the Dreiding force field. With these values, the repeating unit along the chain formed by two CH_2 units is c = 2.5850Å. Note that the planes containing hydrogen atoms at different heights are thus c/2 = 1.2925Å apart. These values are in good agreement with the values of 1.57 Å, 1.10 Å, 108º, and 109º obtained by Bunn in 1939 in an x-ray diffraction study of very long alkyl chains with about 1000 CH_2 units each [11].

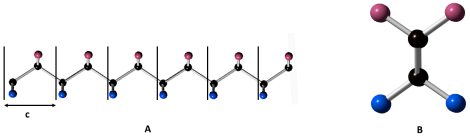

Figure 3. Segment of an infinite aliphatic chain represented in a view perpendicular to the chain direction (A) and in a view along the chain direction (B). Carbon atoms are represented as black balls, while hydrogen atoms are represented by either red or blue balls to highlight their different heights along the chain direction.

6. 2. INTERACTION BETWEEN NEIGHBORING CHAINS

There are different ways in which chains can interact when packed in a crystal. To visualize the different packing arrangements that two neighboring chains may show, it is important to keep in mind that for an efficient packing, the hydrogen atoms from one of the two chains must fit into the pockets formed by the hydrogen atoms of the adjacent chain. Thus, a close packing occurs when the space between molecules is minimized. According to Kitaigorodskii [2], if we consider two chains related by a translation *t* and/or rotated by different angles around their central axis, it is possible to define three basic packing modes. We will name them as M, T, and R packings, with their respective t_M, t_T, and t_R translation vectors (Fig. 4). In the case of the M packing mode, the position of the second chain is obtained from that of the first one by a simple translation, without any rotation relative to the initial chain, in such a way that the two chains have the same relative orientations, with the planes containing the CCC zigzag in a parallel disposition.

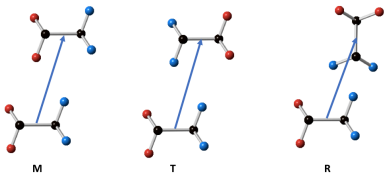

Figure 4. Relative positions of two neighboring chains in the M, T, and R packing modes.

The T packing is obtained by a translation as in the M case, but with an additional 180º rotation around the chain axis, so that the orientation of the CCC planes remains the same as in the M case, but with the positions of the hydrogen atoms exchanged, as it can be seen in figure 4 where the difference between the M and T modes is clearly visible from the different coloring of the H

atoms. Finally, the R packing mode is obtained by a translation and an arbitrary rotation of the second chain around its axis, with the case of a perpendicular or nearly perpendicular orientation of the CCC planes of the two chains, as shown in figure X, being the most usual case.

Packing mode	| t | (Å)	ϕ (º)	H···H (Å)	E_r (kJ mol^{-1})
M	4.330	2 x 74.6	2 x 2.667	-5.9
T	4.150	2x 90.0	2x 2.709	-6.7
R	4.429	77.6, 8.2	2.660, 2.732	-5.7

Table 1. Optimal values for the different parameters characterizing the three basic packing modes. | t | corresponds to the distance between the centers axis of the two chains and ϕ is defined as the angle between the C-C projection on a plane perpendicular to the chain axis and the line joining the centers of the two chains. In the M and T modes the value of ϕ is the same for the two chains. H···H are the distances between the two closest hydrogen atoms between chains and E_r is the energy of the packing mode relative to that of two isolated chains with the same fixed geometry.

To obtain the most favorable relative positions of the two chains in the different packing modes explained above, an energy optimization has been made using the Dreiding force field while keeping the geometry of the individual chains fixed. The most relevant parameters obtained for each mode are gathered in table 1.

The relative energy for each configuration has been calculated as the difference between the energy obtained from optimizing the M, T and R structures and twice the energy obtained for a single isolated polyethylene chain, which is found to be 15.15 kJ mol^{-1} per cell with two CH_2 units.

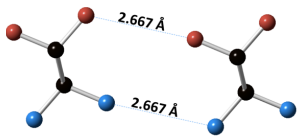

Figure 5. Two polyethylene chains in the most favorable M packing.

For the M packing mode, the minimum energy that two parallel polyethylene chains can have is reached for a value of ϕ = 74.6º (Figure 5). If we compare its relative energy with that for the other packing modes it is found to lead to the second most stable packing, with the shortest distances between hydrogen atoms on different chains of 2.667 Å.

The minimum energy that two antiparallel polyethylene chains (T mode) can have is reached when the two chains are orthogonal to the t_T translation vector (Figure 6). According to our calculations this is the most stable of the three different packing modes, about 0.8 kJmol^{-1} below the optimal M packing.

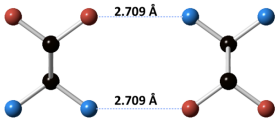

Figure 6. Two polyethylene chains in the most favorable T packing.

In this case, the distances between the hydrogen atoms are somewhat larger, minimizing therefore the repulsion. As a consequence, there will be more space between two chains as in the M case which means a less compact packing when forming crystals.

Since the T mode leads to the absolute minimum of energy, it is not possible to optimize a lowest energy R packing (where the relative rotation of the two chains is arbitrary) since optimization leads always to the T configuration.

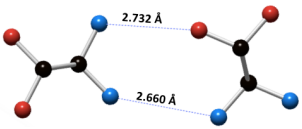

Figure 7. Two polyethylene chains in the most favorable R packing.

To get a representative geometry for this packing mode we have chosen the structure of two polyethylene chains in the experimental crystal structure, which corresponds to an R layout. Although the energy of this configuration does correspond to the absolute minimum of energy it is, however, is a relevant example of this type of packing which is only 0.2 kJmol^{-1} above the most stable M-type packing. In this R configuration the two CCC backbones are nearly perpendicular, with the two ϕ angles close to the 0º and 90º values for a perfect perpendicular orientation. The shortest H···H distances between chains are similar to those in the optimal M mode, allowing thus a similar compact packing in the two modes, both from the geometric and the energetic points of view.

6.3. PACKING OF PARALLEL CHAINS TO FORM LAYERS

Once we have introduced the different ways of packing the dimers, we will focus on the different possibilities to build a layer of parallel chains using the translation vectors of the different packing modes. For the M packing mode, chains can be positioned in a tilted arrangement by an arbitrary angle with respect to line joining the centers of neighboring chains.

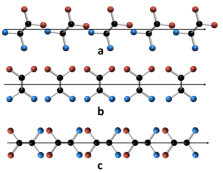

Figure 8. Layers formed by M packing mode.

If we make copies as shown in figure 8a of these chains with the t_M vector along the *a* axis, we will get a plane since the chains are infinite in the direction of CCC axis, which will correspond to the crystal *c* axis.

With an angle $\phi = 74.6º$ we will obtain the most stable layer. Other relevant configurations that can be obtained for the M-type packing are those with the chains perpendicular to the *a* direction (Fig. 8b) or with the chains parallel to it (Fig. 8c). Although from an energetic point of view these arrangements are not as favorable as the one with $\phi = 74.6º$, they may be relevant when building crystals by stacking layers, where the interactions between layers become as important as those within a layer.

For the T packing mode, chains can form a layer as in the previous layout, that is, with their CCC planes inclined at an arbitrary angle with respect to the t_T translation in the *a* direction, with the difference that chains are now antiparallel (Figure 9a).

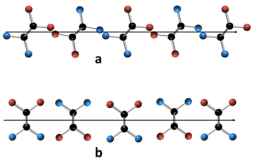

Figure 9. Layers formed by T packing mode.

From the most stable configuration for the dimers (Table 1) we van deduce that the most stable layer will be obtained when the chains are repeated in the T mode with their CCC planes perpendicular to the *a* direction (Figure 9b).

In the case of R packing mode, there are infinite possible arrangements of the chains to form layers and we will only focus on one possibility, alternating chains with perpendicular and parallel orientation to the *a* axis (figure 10). As discussed above, this type of packing (with approximate perpendicular chains) is found in the experimental crystal structure of polyethylene.

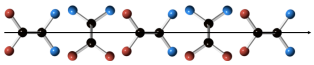
Figure 10. Layer formed by the R packing mode.

It should be noted that there infinitely many possibilities of forming layers using the R packing mode with different orientations for neighboring molecules. If we designate with P and O the chains oriented in a parallel or orthogonal orientation with respect to the *a* direction, the layer shown in figure XX can be designed as ···POPOPOP···, but it is evident that other arrangements such as ···PPOPPOPP··· are possible. These arrangements lead to the so-called *hybrid subcells* when forming crystals through the stacking of planes, which have been found in some cases, although for the sake of simplicity we will not deal with them in this work, which will focus only on the simple crystals with a maximum of two symmetry-independent chains per cell.

Once we have seen how to build layers of chains using the three basic packing modes, our last step in the next section will be to discuss the formation of actual 3D crystals from these layers.

6.4. PACKING OF CHAIN LAYERS TO FORM CRYSTALS

The formation of the simplest crystalline structures for parallel chains arises when considering the different ways of packing layers formed by parallel, perpendicular or tilted chains. Consider, for instance, a layer of tilted chains arranged in the M packing mode as shown in figure 8. The simplest stacking arises from the repetition in the direction perpendicular to the layer (Fig. 8a). Depending on the direction of the inclination of the chains in the planes, it is possible to form alternative structures. One of the most common structure formed by stacking layers of tilted molecules in organic compounds is the so-called herringbone structure. In our case, this structure can be obtained by piling layers of tilted chains with an arbitrary tilting angle ϕ, alternating layers where the chains are tilted in opposite directions, that is ϕ and ϕ+90º, successively, forming the structure that we see in the figure 11.

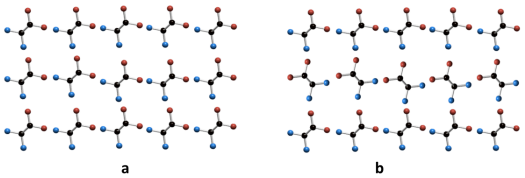

Figure 11. Different structures obtained by piling layers, a) layers of tilted chains have the same arbitrary tiling angle ϕ, b) layers of tilted chains have ϕ and ϕ+90º angle forming herringbone structure.

In 1965 Segerman [12] using symmetry relationships between chains, defined the different modes in which $(CH_2)_n$ chains can be packed in a crystal structure introducing the term *subcell*. Although for infinite chains, these cells are really the unit cells of the crystals containing a $(CH_2)_2$ fragment for each chain, the concept of a subcell is used later in the context of finite chains where it is applied to describe just the packing of the $(CH_2)_n$ fragments, disregarding the end-groups which of course form also part of the complete crystal unit cell. The different subcells described by Segerman are classified in different types by specifying the angles between the zigzag planes of neighbor chains, the type of symmetry features between chains, the angle between the axes and the relative arrangement of these axes.

Considering different symmetry relationships, in reference [13] G. Föster et al. performed energy optimizations for the packing of parallel chains in crystals. If the chains are aligned along the crystal c axis, the different structures are defined by cell parameters a, b and γ, as well as the so-called setting angle τ defined as the angle between the projection of the C-C line on the ab plane

and the crystal a-axis. The different packing modes obtained in this procedure are called basic subcells of the packing. According to Föster et al. we can divide the different subcells into three main groups: orthorhombic, triclinic and monoclinic subcells.

Before delving further into the subject, I would like to highlight the main features of these subcells. In the case of the orthorhombic subcells all angles between cell vectors are 90º with the three edges a, b, c having different lengths. Monoclinic subcells have two angles of 90º and one of them, γ, is different, the three edges also have different lengths. Finally, in the case of triclinic subcells all three angles are different between them and no one has the value of 90º, the edges have different length values between them. To facilitate the comparison between different subcells, however, G. Föster et al. did not use the primitive subcells containing just one chain (Z=1) for the monoclinic and triclinic cases, but instead considered a Z = 2 supercell in which only one of the three cell angles, γ, is different from 90º. The relation between the cell parameters of these supercells and the actual monoclinic or triclinic unit cells can be found in ref. [13].

G Föster based his research on Segerman's results but adapting them to new software and hardware to obtain more accurate results in terms of cell parameters and angle values. Besides this, Förster introduced a new labelling for the different subcells which allowed him to differentiate some cases that were not considered in Segerman's original work. The notation used to designate the subcells in our work follows that of Förster et al. as it allows the explanation of the subcells to be followed more fluently.

The nomenclature uses labels of the type $L\beta_C$ for each structure. All types of subcells begin with an L, this is because the backbone is in the b/b' direction. Subcells with a β are orthorhombic, those with δ are parallel orthorhombic, and subcells that have η and ζ are parallel monoclinic and antiparallel triclinic respectively. Finally, we have a subindex, if we have the letter C or E the structure of the subcell is herringbone, the letter A and G is for pseudoherringbone, and A' or G' for inverse pseudoherringbone structures.

Since for polyethylene, the only case of infinite $(CH_2)_\infty$ chains, only one of the possible structures, the $L\beta_C$ one, has been found, we have decided to obtain the other structures by performing molecular mechanics calculations using the GULP program [4] with the Dreiding force field [9]. In order to generate initial structures for the different subcells given in Förster's paper, we have written a program in Python which uses the geometry of a single chain and approximate cell parameters taken from Förster's paper to generate the coordinates of the 12 atoms (2 x C_2H_4)

contained in each cell. These geometries have been used afterwards as a starting point to optimize the structures shown in figure 12 and to compare their different cell parameters and packing modes using continuous symmetry measures.

All the optimized subcells for the different types of packing are shown in figure 12. Inside the box we can see the orthorhombic subcells with CCC backbones that are not oriented along one of the a, b directions. Subcells $L\beta_C$, $L\beta_A$, $L\beta_{A'}$ on the left side of the box have b > a (with b corresponding to the *y* axis and *a* to the *x* axis). Subcells $L\beta_E$, $L\beta_G$, $L\beta_{G'}$ shown on the right side of the box are those with a >b. If we look at the packing mode of the chains, all orthorhombic cells follow the herringbone structure but with different setting angles. The $L\delta_A$ and $L\delta_G$ subcells are also orthorhombic, but with the difference that they have parallel CCC backbones aligned either with the *a* axis ($L\delta_A$) or with the b axis ($L\delta_G$). The $L\delta_A$ subcell has a setting angle of 0º and the chain occupying the center of the cell is displaced so that it occupies now a position between two chains along the *b* axis. This displacement leads to a T packing along the b direction with consecutive chains alternating their orientation. On the other hand, the $L\delta_G$ subcell has a setting angle of 90º for all chains while the chain of the center is now in an intermediate position, displaced towards the center of the lower edge.

The $L\eta_w$ subcell is triclinic with b > a and the central chain is related to the chain at the origin by a t_T translation, so that it is rotated by 180º around its axis. Both chains are inclined at an angle of 36.4º with respect to the *a* axis. Finally, the $L\zeta_L$ subcell, which is monoclinic, has all the chains with the same inclination and the central chain is related to that on the origin by a t_M translation. Contrary to what happens for the triclinic packing, in this case *a* will be larger than *b*.

Table 2 shows the optimized cell parameters obtained for each subcell, the setting angle τ, the volume per chain (there are two chains per unit cell) and the relative energy per chain taking the experimentally observed $L\beta_C$ structure as a reference.

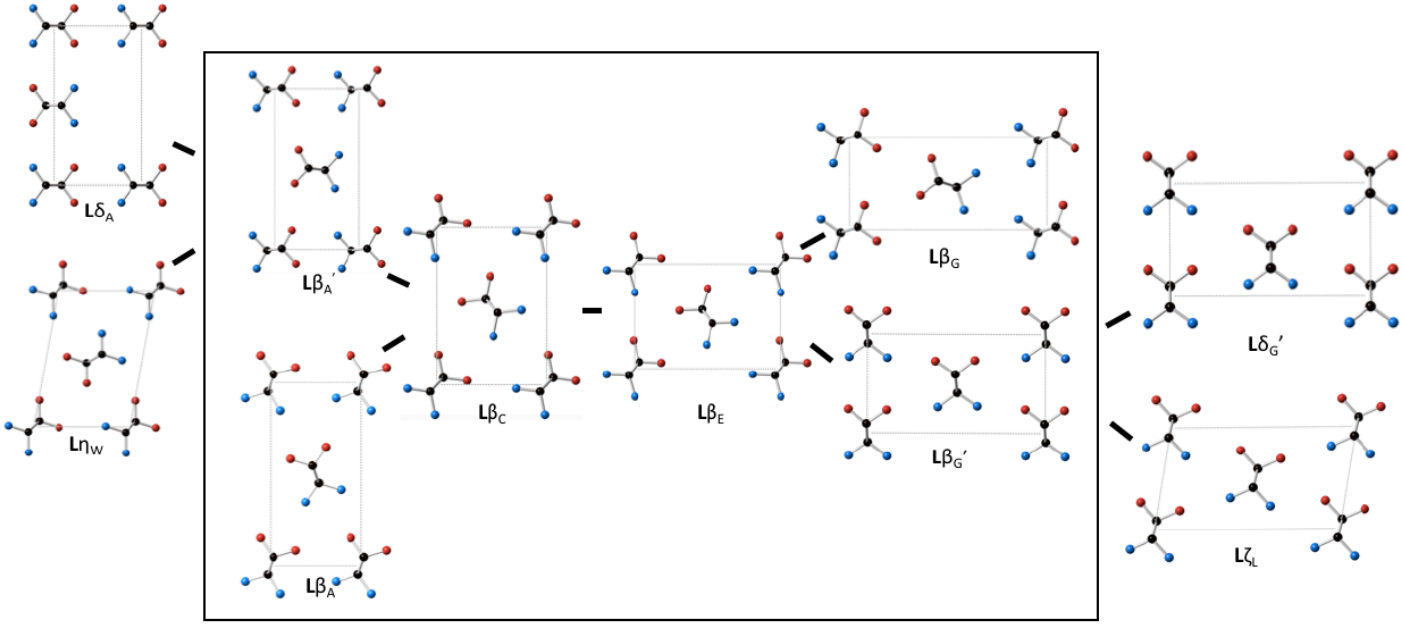

Figure 12. Different subcells obtained from the optimization of the cell parameters for the packing of infinite CH_2 chains with Z = 2 using the Dreiding forcefield.

We have included also the packing factor, that is the percentage of the cell volume occupied by atoms, assuming that they are spherical with van der Waals radii r_H = 1.2Å and r_C = 1.7Å. The packing factor gives a measure of how efficiently both chains are packed in each subcell, with a higher value indicating a more effective packing. The average value for the packing factor for organic crystals is ~ 70%. The last two columns contain the symmetry measures discussed in the next section.

Subcell	a	b	c	γ	τ	V	E_{rel}	PF (%)	S(L_6-HP-6)	S(HPR-12)
$L\beta_C$	5.04	7.28	2.583	90.0	47.1	47.4	0	71.5	0.813	0.753
$L\beta_E$	7.22	5.15	2.582	90.0	51.8	48.0	0.6	70.9	1.097	1.019
$L\beta_A$	4.29	8.63	2.583	90.0	70.8	47.8	0.5	71.0	0.553	0.513
$L\beta_A$'	4.53	8.60	2.583	90.0	7.08	50.3	2	67.2	0.187	0.180
$L\beta_G$	9.17	4.29	2.583	90.0	18.3	50.8	2.2	66.9	1.083	1.020
$L\beta_G$'	8.11	4.52	2.581	90.0	88.8	47.4	0.3	70.8	0.030	0.028
$L\delta_A$	4.52	8.15	2.584	90.0	0.0	47.6	0.2	72.4	5.480	5.101
$L\eta_w$	5.12	7.30	2.582	80.4	36.4	47.9	0.4	71.1	1.623	1.509
$L\delta_G$	8.06	4.53	2.583	89.9	89.1	47.2	0	72.7	0.530	0.491
$L\zeta_L$	8.14	4.55	2.583	80.3	77.5	47.2	-0.1	71.6	0.730	0.676

Table 2. Principal geometric parameters for the different subcells optimized using the Dreiding forcefields. Cell parameters a, b, c in Å; γ and τ in º; volume in Å^3 per C_2H_4 unit. E_{rel} is the relative energy per chain with respect to the experimentally observed LβC structure. PF is the packing factor for each structure. The symmetry measures for the hexagons formed by the points where the central axis of the chains intersects the *a-b* plane and the hexagonal prisms obtained by adding the points where the chain axes cut the plane parallel to the *a-b* plane at a height c. The CSM values obtained by *Cosymlib* are identified by the standard labels in this library: HP-6 for the regular hexagon and HPR-12 for the hexagonal prism, using L_6 for the hexagon formed by the six closest chains. For the case of the hexagonal prism, he have considered as a reference shape a rectangular prism with a regular hexagon at the basis with a side of 4.489 Å (the

average of the six closest distances) and a height c = 2.582 Å, as found in the experimental structure for polyethylene [14]

If we look at the values in the table 2, all subcells have a very similar value for the c parameter, the repeat unit along the chain axis. This means that the different packing modes do not introduce significant changes to the structure of the individual chains when packing them to form a crystal at moderate temperature / pressure conditions. The average values for the CC, CH bond distances and CCC and HCH angles are 1.551 Å, 1.093 Å, 112.71º and 108.15º respectively, very close to those obtained when optimizing the geometry for an isolated chain. In figure 13 we represent the energy per chain as a function of the volume per chain. If we look at the results obtained, we see that as expected, that close packing (smaller volumes) leads to more stable structures (smaller energies), although this trend is not strictly followed by structures with very similar volumes. Surprisingly it can be also seen that the experimental structure does not correspond to the most stable packing mode $L\zeta_L$ which has lower energy and volume values per chain than the $L\beta_C$ structure found for polyethylene [14].

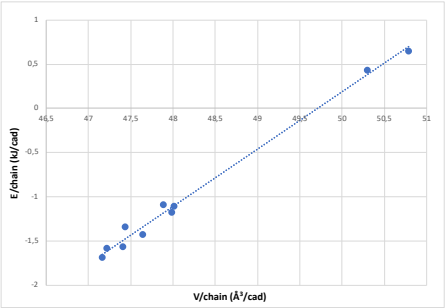

Figure 13. Representation of the energy per chain as a function of the volume per chain for the optimal subcells obtaining with the Dreiding force filed.

It has been argued in the literature [13] that the $L\beta_C$ is slightly favored over the most stable $L\zeta_L$ one if one considers the vibration of the chains in a free energy optimization at ambient pressure and temperature values.

6.5. CONTINUOUS SYMMETRY MEASURES FOR THE DIFFERENT PACKING MODES

If we consider ideal cylindrical chains, the most compact packing for this type of chain would be hexagonal packing, where each chain by wrapped in six neighboring chains at the same distance forming a regular hexagon. The packing factor for hexagonal packing is 0.91, that is 91% of the cell is occupied by atoms and 9% by gaps. But as we've seen in the previous section, none of the packing obtained optimizing the structure is hexagonal showing that in this case hexagon packing is not the most energetically favorable. This is because in this case we don't have exactly cylindrical chains, but polyethylene chains have a lower symmetry section than a circle therefore

the packing that will optimize the occupied space will be less symmetrical with an arrangement of neighboring chains in a more or less distorted hexagonal geometry.

Polyethylene chains, which have an average section with rectangle symmetry, have a C_2 symmetry axis in the direction of the chain, therefore, the hexagon formed by the six nearest chains is expected to maintain at least one C_2 axis in the perpendicular direction.

This fact would give us three possible types of distortion (figure 14): a lower symmetry distortion where three pairs of side are not equal (figure 14a), another distortion where we have four equal short sides and two longer equals (figure 14b) or the opposite with four equivalent long ones and two shorter and equal sides (figure 14c).

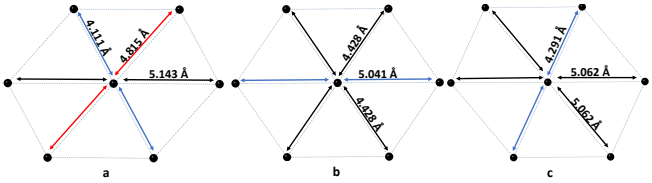

Figure 14. Different types of distortion in a hexagon with the different values of the sides.

Another type of distortion to consider is due to a change in the position of the central chain, that both in the regular hexagon and in all the three distortions discussed above occupies the geometric center of the hexagon. As we have observed in the figure 12, in the $L\delta_A$ and $L\delta_G$ subcells the central chain is displaced from the geometric center. In the case of $L\delta_A$ subcell, this displacement is extreme, giving rise to a packing derived from that of a square packing, which cannot be considered as a distortion of a hexagonal packing. In these distortions where the central chain does not occupy the geometric center, it would break the axis of symmetry C_2, if we consider the geometric object formed by the hexagon of the neighboring chains and the central chain.

Apart from the hexagons corresponding to the section of the packing of the chains a plane perpendicular to the direction of the chains to be able to make a detailed study of the packing of chains, the hexagonal prisms corresponding to each subcell have been considered taking a height c, that is the height corresponding to the elementary unit C_2H_4 of which the chains are formed. In the case of infinite chains, these hexagonal prisms do not provide much more information than the corresponding hexagons, but in the case of finite chains where we can have an additional distortion consists of the inclination of the chains respect to the a-b plane of the cell. We will see that the prisms are more useful to compare the packing in different compounds. Therefore, we

will also include here the measurements for the hexagonal prism that will be needed later to compare with calculations obtained for structures with finite chains.

6.6. EXPERIMENTAL STRUCTURES FOR POLYETHYLENE

In this section we will discuss the structure of the polyethylene obtained experimentally. In 1939 Bunn [11] wanted to study crystals of long paraffin chains using X-ray crystallographic methods to see what changes they might have in their structure. In his study he included paraffin chains with approximately 1000 atoms as a limiting case where the chains are so long that the interactions of the terminal groups are insignificant. Therefore, this structure is taken to be equivalent to that of polyethylene since the predominant interactions are those of the CH_2 groups of the chain. In 1975 the structure of polyethylene was studied again but, in this case, it was done at low temperatures (4K and 90K) [14] using neutron diffraction and potential energy calculations for deuterated polyethylene. The samples were prepared by polymerization of deuterated ethylene using Ziegler-Natta catalysts.

Polyethylene is found to crystalize in the orthorhombic Pnma group with two C_2H_4 chains per unit cell (Z = 2) and a = 4.851 Å, b = 7.121Å, and c = 2.548Å at 4K.This structure coincides with the $L_\beta C$ subcell in the article by Föster et al. and has the chains tilted by an angle ϕ and the central chain is rotated ϕ + 90º with ϕ = 41º at 4K. The values for the S(L_6-HP-6) and S(HPR-12) measures for the experimental structures of polyethylene are 0.677 and 0.656 for the hexagon and 0.629 and 0.609 for the prism at 4 and 90K, respectively. These values compare well with those obtained for the optimized structure, although they are somewhat smaller than the computed ones. The cell constant and atomic coordinates are very similar at the two temperatures. Accurate positions for the atoms were obtained by fitting the crystal structure using theoretical calculations to minimize the potential energy as a function of the CC, CH bond lengths and CCC and HCH the angles. The results obtained by Bunn are very similar to those obtained from the $L_\beta C$ subcell.

7. PACKING OF FINITE ALIPHATIC $(CH_2)_N$ CHAINS

In this section we will discuss the packing of finite aliphatic chains with a number of carbon atoms between 6 and 60. Chains with an odd number of carbons have been seen to have a different behavior when forming crystals than those with an even number of carbon atoms, but first we will explain the principal features of a single chain. Then we will continue with the packing of parallel chains to form layers, and finally the stacking of these layers to form crystals. In the last section we will discuss the features of the experimental structures that have been found using the Cambridge database and compare their packings with those found for infinite chains in the previous section.

7.1. STRUCTURE OF A SINGLE CHAIN

The structure that finite aliphatic chains adopt in crystals is basically the same as that described above for infinite chains, with the basic difference that in finite chains we have the terminal groups at both ends, for instance, CH_3 groups in the simplest case of linear alkanes. The chain adopts an all-trans configuration with its characteristic CCC zigzag backbone with all carbon atoms on the same plane and the hydrogen atoms of the CH_2 groups alternating from one side of the CCC backbone to the other. In simple paraffins, two of the hydrogen atoms of the CH_3 terminal groups lie on one side of the backbone as those of the central CH_2 groups, while the third atom, marked in green in the figure lies on the plane containing all carbon atoms.

Figure 15. Finite aliphatic chains with nine (left) and ten (right) carbon atoms, C_9H_{20} and $C_{10}H_{22}$.

The main structural difference between alkane chains with odd or even number of C atoms is in their symmetry. While the plane containing the CCC backbone is a symmetry plane for both cases, for chains with an odd number of carbon atoms we have an additional perpendicular symmetry plane containing the central CH_2 group and a C_2 axis at the intersection of the two planes, giving a C_{2v} symmetry for the whole structure. In the case of chains with an even number of C atoms we find a center of inversion at the middle of the central C-C bond and a C_2 axis perpendicular to the CCC backbone passing through the inversion point, resulting in a C_{2h} symmetry. As a consequence of this different symmetry, in the odd carbon atom chain in figure 16, we can see that hydrogen atoms corresponding to the terminal groups perpendicular to the CCC backbone

are on the same side of the chain, while in the case of an even carbon chain they are at opposite sides of the chain. The same behavior is found for H atoms on the CCC plane: for the chain with an odd number of C atoms they lie at the same side of the line joining all centers of the C-C bonds, while for chains with even number of carbons they lie at opposite sides of this line.

To study the variation of the length of chains with the number of CH_2 groups, a graphical representation has been made (figure 16). If we do a regression, we obtain l = 2.8499 + 1.2768n where l is the length of the chain and n the number of CH_2 groups. From this equation we can deduce that the average length of a CH_2 group has a value of 1.2768Å and for the CH_3 group it would be 2.8499/2 since there is one at each end of the chain. The obtained value of c = 2.5536Å is very similar to that of infinite chains.

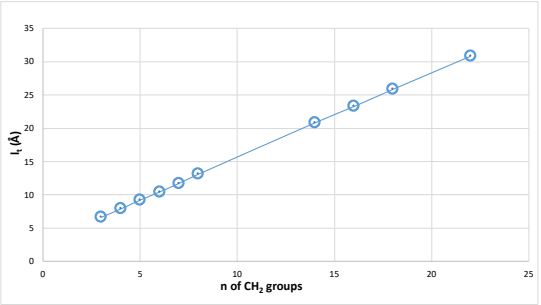

Figure 16. Representation of the length of chains as a function of the number of CH_2 groups.

These results are very similar to those obtained by V. Chevalier et al. [15] who carried out a study of the number of carbons correlating it with the cell parameter c.

7.2. INTERACTION BETWEEN NEIGHBORING CHAINS

When we talk about finite aliphatic chains with a large number of carbons, the interactions between neighboring chains are similar to those in infinite aliphatic chains, since the effect of the interaction between terminal groups becomes negligible when the chain has a considerable number CH_2 groups. This is no longer true when we study shorter chains, where the interactions between terminal groups can produce alterations in the crystal structure. To take this into account, Müller [16] interpreted that the energy of the corresponding packing was composed of two components, one associated with the lateral interactions of the molecular chains and the other with the interaction of the terminal groups. In general, he observed that when the transition to smaller molecular chains occurred, the proportion between the *a* and *b* cell parameters changed and significant changes occurred in the crystal structure depending on the nature of the terminal groups. We will focus our attention here to the simplest case of normal alkanes, where we have CH_3 methyl groups capping the ends of the chains. For this case, where the nature of the terminal

groups is very similar to the CH_2 groups forming the chain, the effect of terminal groups in the lateral packing of the chains can be neglected for chains with more than 5 carbon atoms.

However, even though in finite aliphatic chains we have to consider the effects of terminal groups, when we deal with the different ways of packing chains together, they have the three same basic packing modes, M, T, and R as for infinite aliphatic chains, although for the case of finite chains we must consider the possibility that the translation vectors might have a component in the direction of the chain axis. To visualize this, let us take the example of the shortest chain in our study, C_6H_{14}, although the general concepts explained here will apply for any finite chain, irrespective of its carbon atom number being even or odd.

In figure 17 we can see a dimer formed by two C_6H_{14} chains related by a simple translation as in the M-packing mode in both a view in a direction perpendicular to the chain axis and in a view along the chain direction

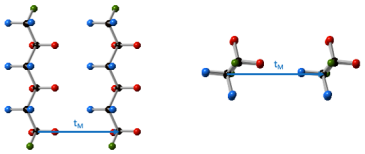

Figure 17. Dimer of C_6H_{14} chains formed by applying a translation in a plane perpendicular to the chain axis (M-packing mode). Lateral view (left) and view along the chain axis (right).

Note that when applying a translation in a direction perpendicular to the chain axis, the terminal groups of the two chains remain at the same height, a situation that might be unfavorable if the interaction between those groups is repulsive. A possible way of minimizing these repulsions it to allow a relative slipping of the two chains in the direction of their axes, a situation shown in figure 18. In the case of linear alkanes, where the size of the terminal CH_3 groups is similar to that of the CH_2 ones, the optimal displacement is of just the length corresponding to one CH_2 group, that is, approximately 1.27Å in the direction of the chain axes. Note that the displacement along the chain by the length of just one CH_2 group results in a switch of the height of the H atoms along the $(CH_2)_n$ backbone which is clearly seen in the figure 18 where the H atoms have been colored according to their position along the chain direction. In this respect, if we disregard the terminal groups, the effect of a translation in the direction perpendicular to the chains plus a slippage by one CH_2 unit along the chain direction is equivalent to the T-type packing mode for infinite chains. It is, however, important to notice that the effect has been achieved without a rotation around the chain axis at it is evident from the top view in figure 18 where the color of the H in the CH_2 groups

has been reversed, but the orientation of the terminal H atom, marked in green, is the same in the two chains and not rotated by 180º.

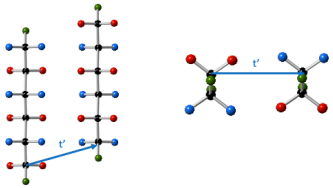

Figure 18. Dimer of C_6H_{14} chains formed by applying a translation with a c/2 displacement in the direction of the chains. Lateral view (left) and view along the chain axis (right).

7.3. PACKING OF PARALLEL CHAINS TO FORM LAYERS

When considering the packing of parallel chains to form rows we have the same possibilities as for infinite chains, with the additional complication that now we might consider the relative slippage of the chains along the direction of their axes. A row of chains related by a t_R translation as found in the orthorhombic polymorph of chains with an even number of carbon atoms is shown in figure 21b. These rows of chains can further stack in a direction perpendicular to the chains to form a layer with a finite thickness given by the length of the chains. As we will discuss in the next section there will be different possibilities when piling these layers to form a crystal, although these possibilities depend only on the interactions between terminal groups of successive layers which are not the main interest in the present work.

A more interesting question is the effect of the relative slippage of the chains along their axes in the formation of these layers. In figure 19 we can see a row of chains where each consecutive chain is displaced by the height of one CH_2 unit in the direction of the chains. If we represent it with the chains aligned with the vertical axis as in the left part of the figure we can clearly see that the position of the same group in each chain, for instance the terminal methyl group at the bottom of the chains is progressively displaced in the vertical direction. An alternative representation of the same row is to keep all equivalent atoms on each chain at the same height as in the right part of figure 19. In this representation it is more obvious that the chain axes are tilted by an angle α with respect to the translation vector generating the row. Since the optimal length of the translation vector (t ~ 4.11 Å) depends basically on the lateral $CH_2 \cdots CH_2$ interactions and the length along the chain of a single CH_2 group (d ~ 1.27 Å) is practically the same for any chain and $\cos \alpha = d/t$, a value of $\alpha \sim 72º$ is found for all phases with this packing mode.

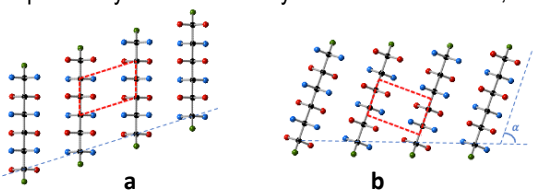

Figure 19. Row of C_6H_{14} chains formed by applying a translation in a direction

perpendicular to the chain axis. a) View with the chains aligned in the vertical direction and b) view with all equivalent atoms at the same height in the vertical direction. The red dashed lines show the subcell used for describing just the central part of the chains formed by the CH_2 groups in such a way that it can be directly compared to the crystallographic cells for infinite chains depicted in figure 12.

In order to compare the packing of chains in the different polymorphs it is useful to use a cell with the chains aligned along the *c* direction with a C_2H_4 fragment per chain where *c* is perpendicular to the *a-b* plane as for the case of the infinite chains. Since this cell, shown in red in the right part of figure does not coincide with the real unit cells of the crystals, which contain the whole chain, including the terminal groups at both ends, the term *subcell* is often used to design it note that if we try an analogous definition of the subcell using the orientation depicted in the left part of the figure, we would obtain a cell with the chain axis in the c direction, but tilted with respect to the a-b plane, and this would complicate the comparison with the cells obtained for the infinite chains.

Another important result that is a direct consequence of the tilting of the layers is that the thickness of the layers is not only dependent on the length of the chains l, but also on their inclination given by α. Since $\alpha \sim 72º$ is found in all cases, the layer thickness, $l \cdot \sin \alpha$ is about 0.95 l, where l is a linear function of the number of CH_2 groups as explained above.

As in the case of infinite chains, there are different possibilities to form layers by stacking rows of chains as those explained above. Although basic crystallographic information is available for a number of n-alkanes, complete single-crystal analysis has been limited to only a few cases which will be described in the next section. This has been mainly due to problems in growing large single crystals of sufficient size, quality and purity suitable for X-ray and neutron diffraction methods. Three solid-state structures have been identified: triclinic, monoclinic, and orthorhombic.

In the triclinic phases, rows of chains such as that shown in figure 19 are propagated in a direction perpendicular to the chains. The resulting layer is shown in figure 20. In the left part of the figure the chains are seen to be tilted by $\alpha \sim 72º$ with respect to the planes containing equivalent atoms, while in the view along the direction of the chains in the right part of the figure it is clearly visible that the subcell in this layers is of the triclinic $L\eta_{w.}$ type.

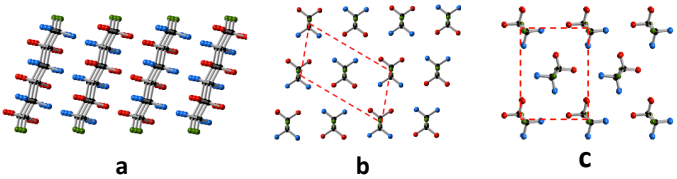

Figure 20. Layer of C_6H_{14} chains in the triclinic polymorph in a) a lateral view and b) a view along the direction of the chains. The red dashed lines indicate the $L\eta_w$ type subcell in these layers. c) the orthorhombic polymorph in a view along the direction of the chains. The red dashed lines indicate the $L\beta_C$ type subcell in these layers.

The subcell in the monoclinic and orthorhombic polymorphs is in both cases of the $L\beta_C$ type with the typical herringbone pattern, although there is a fundamental difference between both cases. While in the monoclinic polymorph the chains are tilted by an angle of ~ 72º with respect to the planes containing equivalent atoms as for the triclinic phase, in the orthorhombic polymorphs they are all oriented perpendicular to the planes containing equivalent atoms.

7.4. PACKING OF CHAIN LAYERS TO FORM CRYSTALS

The formation of full crystalline structures results from the stacking of layers formed by parallel chains, which may be either perpendicular or tilted with respect to the planes containing equivalent atoms as shown in figure 21 for the triclinic and the orthorhombic crystals of n-hexane. Since we are mainly interested in the lateral packing of long aliphatic chains and the different ways in which the layers may be stacked depends only on the interactions between the terminal groups of the molecules in opposite layers, we will not discuss in detail the different possibilities observed experimentally, which lead to cells pertaining to different space groups.

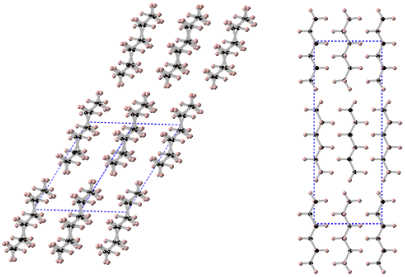

Figure 21. Stacking of three layers of C_6H_{14} chains in a) the triclinic and b) the orthorhombic polymorph in a view perpendicular to the chain axes. The blue dashed lines indicate the unit cell of the two crystals, with one and four molecules per unit cell in the triclinic and the orthorhombic cell, respectively.

A further particularity of the n-alkanes is that of the so-called rotator phases. In the case of finite aliphatic chains, as we increase the temperature and we approach the melting temperature, the chains are able to rotate around their axes, changing their orientation with respect to their neighbors.

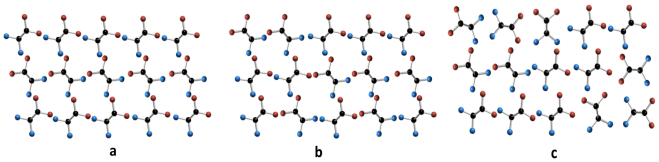

Figure 22. Different structures formed from the stacking of chain layers a) shows the herringbone structure b) the R_I rotator phase and c) the hexagonal R_{II} rotator phase.

As we can see in figure 22a, the orientation of the chains in the layers of orthorhombic crystals is by default the same as when we were talking about infinite aliphatic chains, the herringbone structure with a $L\beta_C$ type subcell. But as we increase the temperature, we can have five rotator phases where some of the chains have changed their orientation with respect to that of their neighbors. The five rotator phases have all a three-dimensional crystalline positional order for the position of the chains and the molecules are arranged in layers such as those discussed above. However, as the temperature increases the rotational order decreases, and this allows more freedom to the chains to rotate around their axes taking different orientations that do not occur in the crystalline phases at lower temperatures as shown in figures 22b and 22c. At sufficient high temperatures the orientation of neighboring chains is totally random leading to the R_{II} hexagonal phase, where on average, the interactions between one chain and its six closest neighbors are equal, a situation in which the average shape of the chains may be considered to be cylindrical.

7.5. CONTINUOUS SYMMETRY MEASURES FOR EXPERIMENTAL STRUCTURES

In a search in the Cambridge Structural Database (CSD) for long-chain alkanes C_nH_{2n+2} with $n \leq 30$ we found only nine complete structures with cell parameters and atomic positions, seven of them with an even number of carbon atoms and only two with an odd number. Besides this, we have precise determinations of cell parameters using synchrotron radiation [17] for chains with up to 60 carbon atoms, many of them with orthorhombic structures for which it is easy to determine the position of the six closest neighbor chains without knowing the exact atomic positions. A third source of information is the study by Doucet et al. [18] where the evolution of cell parameters for the orthorhombic structures of C_nH_{2n+2} with n = 17, 19, and 20 as a function of the temperature is reported. In this case we are interested in showing how the hexagonal continuous symmetry measure is able to nicely follow the evolution from an orthorhombic to hexagonal phase.

Let us start our discussion with the structures retrieved from the CSD. All retrieved structures correspond to the triclinic polymorph in which we can find layers of tilted molecules with an

inclination angle α ~ 72º. As we have previously explained, the chains are not totally cylindrical and therefore they assemble in a distorted hexagonal packing. In the studied structures the distortion we found corresponds to a hexagon with four long sides and two short sides. Since the chains are tilted, we have examined both the hexagon in the plane containing equivalent atoms and the hexagon in a plane perpendicular to the chain direction. These values can be found in the table 2A in the appendix. To be able to compare the differences in the packing for finite and infinite chains we have computed also shape measures for the hexagonal prism (both the oblique and straight versions) with a height corresponding to two CH_2 groups to check if the increase of carbon atoms in the chain affects the symmetry of the packing.

As discussed in the previous section, in the triclinic polymorph we find a $L\eta_w$ type subcell for which we found S(L_6-HP-6) = 1.623 and S(HPR-12) = 1.509. The values found for finite chains are, in average, somewhat larger, 2.149 and 2.699, respectively. But if we disregard the data corresponding to $C_{20}H_{42}$, we find a clear trend of decreasing values as the chain length is increased with the values for n = 24 closely approaching the limiting value for the infinite chains.

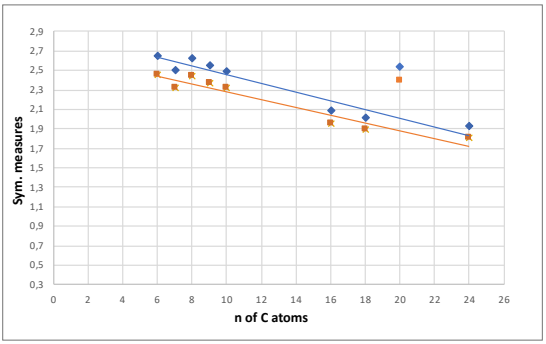

Figure 23. Representation of the symmetry measures S(L_6-HP-6) and S(HPR-12) of triclinic structures. The measures for the hexagon are those of blue color and those of orange belong to the hexagonal prism.

A linear regression has been performed for the hexagon and hexagonal prism measures. The equation obtained for S(L6-HP-6) is S = -0.0443n + 2.8985 with R_2 = 0.9229 and for S(HPR-12) S = -0.0398n + 2.6795 with R_2 = 0.9222.

Although it would be interesting to have further values to confirm this, it has been found that for longer chains the monoclinic or orthorhombic phases become the most stable polymorphs and no more data for triclinic phases is available.

Since it is not necessary to dispose of accurate atomic positions to calculate the symmetry measures for the chain arrangement in the orthorhombic polymorph, we include CSM values for this type of packing in the table 3A in the Appendix. In the orthorhombic structure the chains are perpendicular to the planes containing equivalent atoms (figure 21b) and the prism with the long

edges parallel to the chain direction is already a straight prism which can be directly compared with those obtained for the infinite chains. As discussed above, the packing in the orthorhombic phase presents the typical herringbone arrangement corresponding to the Lβ_C type subcell for which S(L_6-HP-6) = 0.813 and S(HPR-12) = 0.753. The values gathered in the table 3A in the appendix lie in the range 0.6 – 0.4 with average values 0.526 for the hexagon and 0.488 for the prism, somewhat lower than the computed values for the the Lβ_C type subcell. Remember, however that values obtained for the experimental structure of polyethylene were also somewhat smaller than the computed ones so that the Dreiding forcefield which seems to slightly overestimate these parameters. Note, however, that the CSM obtained for the orthorhombic structures are clearly lower than those obtained for the triclinic phases indicating a more symmetric lateral packing in the layers of the orthorhombic crystals. Although the values for the shorter chains seem to be somewhat larger, a plot of the CSM as a function of the chain length (figure 24) does not indicate any clear trend as that found for the triclinic phases where the decrease in the CSM values with the chain length is clearly evident.

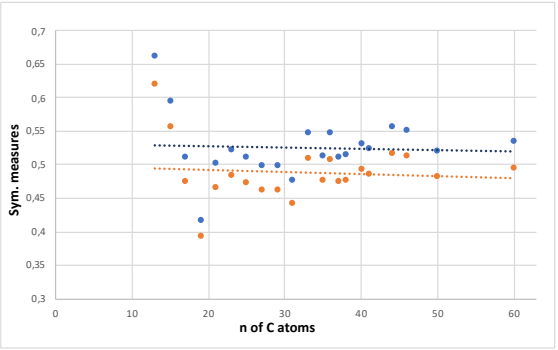

Figure 24. Representation of the symmetry measures S(L_6-HP-6) and S(HPR-12) of orthorhombic structures. The measures for the hexagon are those of blue color and those of orange belong to the hexagonal prism.

The last point we would like to address is the utility of the CSM in the characterization of the structural changes induced in the packing of the layers by an increase of the temperature. In an article by J. Doucet et al. [19] the evolution of the cell parameters *a*, *b* and *c* of the chains $C_{17}H_{36}$, $C_{19}H_{40}$ and $C_{21}H_{44}$ was studied as a function of temperature. Using a Python program written by us the points corresponding to the hexagon for those cell parameters have been calculated so that it can be then compared with the regular hexagon using *Cosymlib*. The data obtained for the hexagonal symmetry measure have been represented as a function of temperature in figure 25.

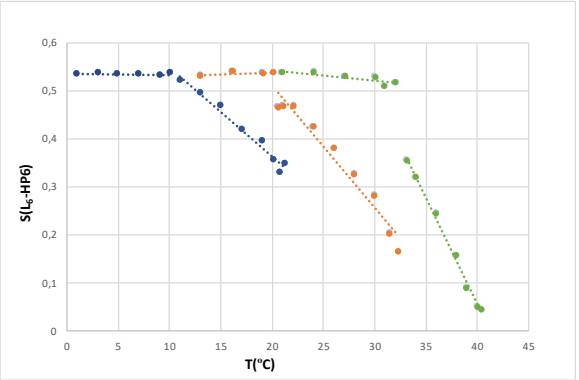

Figure 25. Hexagonal symmetry measure S(L_6-HP6) for the packing of neighbor chains in C_nH_{2n+2} with n = 17, 19, and 21. Data in blue correspond to $C_{17}H_{36}$, in orange to to $C_{19}H_{40}$, and in green to $C_{21}H_{44}$.

In figure 25 it can be clearly seen that at low temperatures, all three phases have on orthorhombic structure with temperature independent S(L_6-HP-6) values in the range 0.5 – 0.6 with an abrupt change at the transition temperature to the rotator phase for which a linear decrease of S(L_6-HP-6) with temperature is observed with values closer to 0 which indicates that the structure becomes progressively more symmetrical, closer to the regular hexagon. The equations obtained from a linear fit to the data at high temperatures are $S = -0.0187\ T + 0.7367$ with $R^2 = 0.9761$ for $C_{17}H_{36}$, $S = -0.0252\ T + 1.0125$ with $R^2 = 0.959$ for $C_{19}H_{40}$, and $S = -0.0438\ T + 1.8016$ with $R^2 = 0.995$ for $C_{21}H_{44}$ indicating that the rate of the symmetrization of the packing in the layers as a function of the temperature is clearly increased as the chain becomes longer, with the $C_{21}H_{44}$ layers reaching an almost perfect hexagonal symmetry just by increasing the temperature about 10ºC.

10. CONCLUSIONS

In order to study the crystalline packing in molecules with long aliphatic chains, in our case the n-alkanes, we use continuous measures of symmetry to carry it out. To be able to compare the data obtained for experimental structures with the theoretical data for ideal infinite chains, different structures have been optimized for polyethylene using molecular mechanics methods. From these calculations, we may conclude that:

1) Polyethylene chains can be packed to form crystals in different ways with very similar energies and packing factors.
2) There is a general trend according to which the more compact the packing, the more stable the structure, although there may be changes in this trend for very similar volumes.
3) The symmetry measures obtained for the hexagons/ hexagonal prisms from the six closest chains allow us to differentiate the different types of packing, especially between the Lβ_C and Lη_w subcells, which are those found experimentally for finite alkane chains.
4) We found a very favorable packing, $L\delta_A$, that is not derived from the hexagonal packing but from the square one. The symmetry measures for this packing mode differ considerably from those that come from the hexagonal close packing. The $L\delta_A$, packing is quite compact and stable since the perpendicular section of polyethylene chains has a rectangular symmetry which is well suited to compactly fill the space in a lattice derived from the square packing.

Once the calculations for the infinite chains have been carried out, I have applied these concepts to study the packing for finite chains obtained from the data base or from the literature. The conclusions that can be drawn are the following:

1) In the triclinic phases the chains pack with a Lη_w-type subcell and the shape measures for the perpendicular prisms are similar to those obtained for polyethylene with the Lη_w

structure, although slightly larger than those found for infinite chains. The symmetry measures decrease, however, with increasing number of carbon atoms in the chains.

2) Shape measures for the normal and the perpendicular prisms have a difference of about 0.5 for all cases. This indicates that the distortion due to the tilting of the chains is the same for all compounds, in agreement that the inclination angle is approximately 72º for all of them, independent of the chain length.
3) The data found for $C_{20}H_{42}$ do not follow the trends observed in other structures, hinting to some problems in the experimental structure resolution.
4) Orthorhombic phases have a packing of the Lβ_C type and the symmetry measures smaller than for triclinic phases. We can conclude, thus, that packing in the orthorhombic phases is more symmetrical (closet to hexagonal) than in the triclinic ones.
5) The values of the measures as a function of the number of carbon atoms in the chain do not show large variations and no obvious trend can been found.

To finish, for the rotating phases in $C_{17}H_{36}$, $C_{19}H_{40}$, and $C_{21}H_{44}$ we can conclude that:

1) By the representing the symmetry measures as a function of the temperature, we can clearly detect the orthorhombic to R_I phase transition.
2) In the orthorhombic phase, symmetry measures do not change in the studied temperature interval, with values between 0.5 - 0.6 for all three compounds.
3) For the rotator phases we find a linear decrease in the symmetry measures as a function of the temperature with a slope that increases as the chain is longer, that is, the longer the chain the faster the symmetrization of the layers when increasing the temperature.
4) For $C_{21}H_{44}$ in 10ºC from the transition the hexagonal symmetry has already been practically reached.

11. REFERENCES AND NOTES

[1] Kitaigorodsky, A. I. Molecular crystals and Molecules. (Academic Press, 1973).
[2] Kitaigorodsky, A. I. Organic Chemical Crystallography. (Consultants Bureau, 1961).
[3] Groom, C. R. & Allen, F. H. (2014). Angew. Chem. Int. Ed. 53, 662-671. [Older reference: Allen, F. R. (2002). Acta Cryst. B58, 380-388.]
[4] Gale, J. D. GULP: A computer program for the symmetry-adapted simulation of solids. J. Chem. Soc., Faraday Trans., 93, 629–637 (1997).
[5] Palmer, D. C. Visualization and analysis of crystal structures using CrystalMaker software. Zeitschrift für Kristallographie - Crystalline Materials vol. 230 (2015).
[6] Gaussian 09, Revision 0.01, M. J. Frisch et al., Gaussian, Inc., Wallingford CT, 2009.
[7] (a) Allinger, N. L. J. Am. Chem. Soc. 1977, 99, 8127. (b) Sprague, J. T.; Tai, J. C; Yuh, Y.; Allinger, N. L. J. Comput. Chem. 1987,8,581. (c) Burkert, U.; Allinger, N. L.; Molecular Mechanics·, American Chemical Society: Washington, DC, 1982; ACS Monograph 177.
[8] (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ohio, C.; Algona, G.; Profeta Jr., S.; Weiner, P. J. Am. Chem. Soc. 1984,106, 765. (b) Weiner, S. J.; Kollmann,
[9] Mayo, S. L., Olafson, B. D. & Goddard, W. A. DREIDING: a generic force field for molecular simulations. J. Phys. Chem., 94, 8897–8909 (1990).
[10] Carreras, A., Bernuz, E., Marugan, X., Llunell, M. & Alemany, P. Effects of Temperature on the Shape and Symmetry of Molecules and Solids. Chem. Eur. J 25, 673–691 (2019).
[11] Bunn, C. W. The crystal structure of long-chain normal paraffin hydrocarbons. The 'shape' of the $>CH_2$ group. Trans. Faraday Soc., 35, 482–491 (1939).
[12] Segerman, E. The modes of hydrocarbon chain packing. Acta Cryst., 19, 789–796 (1965).
[13] Förster, G., Meister, A. & Blume, A. Chain packing modes in crystalline surfactant and lipid bilayers.Curr. Opin. Colloid Interface Sci., 6, 294–302 (2001).
[14] Avitabile, G. et al. Low temperature crystal structure of polyethylene: Results from a neutron diffraction study and from potential energy calculations. J. Polymer Sci. Polymer Lett. Ed, 13, 351–355 (1975).
[15] Chevallier, V., Petitjean, D., Ruffier-Meray, V. & Dirand, M. Correlations between the crystalline long c -parameter and the number of carbon atoms of pure n -alkanes. Polymer, 40, 5953–5956 (1999).
[16] A. Müller, Proc. R. Soc. London Ser. A 138, 514 (1932).
[17] Craig, S. R., Hastie, G. P., Roberts, K. J. & Sherwood, J. N. Investigation into the structures of some normal alkanes within the homologous series C13H28 to C60H122 using high-resolution synchrotron X-ray powder diffraction. J. Mater. Chem. 4, 977 (1994).
[18] Doucet, J., Denicolo, I. & Craievich, A. X-ray study of the '"rotator"' phase of the odd-numbered paraffins C17H36, C19H40, and C21H44. J. Chem. Phys. vol. 75 1523–1529 (1981).
[19] Denicolò, I., Doucet, J. & Craievich, A. F. X-ray study of the rotator phase of paraffins (III): Even numbered paraffins C18H38, C20H42, C22H46, C24H50, and C26H54. J. Chem. Phys., 78, 1465–1469 (1983).

APPENDICES

APPENDIX 1: TABLES OF VALUES

This appendix shows the values obtained from the symmetry measures for infinite and finite chains and the distances obtained from the center to the vertex

Subcells	d_1(Å)	d_2(Å)	S(ML_6-HP-6)	S(L_6-HP-6)	S(HPR-12)
$L\beta_C$	2x 5.041	4x 4.428	0.813	0.813	0.753
$L\beta_E$	2x 5.150	4x 4.434	1.097	1.097	1.019
$L\beta_A$	4x 4.822	2x 4.293	0.553	0.553	0.513
$L\beta_A$'	4x 4.852	2x 4.541	0.187	0.187	0.180
$L\beta_G$	4x 5.062	2x 4.291	1.083	1.083	1.020
$L\beta_G$'	4x 4.645	2x 4.525	0.030	0.030	0.028
$L\delta_A$*	2x 4.074, 2x 4.526	4x 6.089	6.556	5.480	5.101
$L\eta_w$	2x 4.815, 2x 5.143	2x 4.111	1.623	1.623	1.509
$L\delta_G$	2x 4.535, 2x 5.000	2x 4.325	0.676	0.530	0.491
$L\zeta_L$	2x 4.984, 2x 4.551	2x 4.318	0.730	0.730	0.676

Table 1A. Geometric parameters characterizing the packing of chains in the different subcells including the two shortest distances between a given chain and its neighbors (d1 and d2) and the symmetry measures for the hexagons formed by the points where the central axis of the chains intersects the a-b plane and the hexagonal prisms obtained by adding the points where the chain axes cut the plane parallel to the a-b plane at a height c.

Chain	d_1(Å)	d_2(Å)	S(N-L_6-HP-6)	S(N-HPR-12)	$S_{\perp}$ (L_6-HP-6)	$S_{\perp}$(HPR-12)
C_6H_{14}	2x 5.403, 2x 4.696	2x 4.131	2.448	2.982	2.648	2.454
C_7H_{16}	2x 5.356, 2x 4.690	2x 4.150	2.205	2.730	2.496	2.315
C_8H_{18}	2x 5.387, 2x 4.686	2x 4.123	2.429	2.954	2.628	2.438
C_9H_{20}	2x 5.341, 2x 4.674	2x 4.118	2.285	2.783	2.549	2.365
$C_{10}H_{22}$	2x 5.413, 2x 4.724	2x 4.174	2.296	2.833	2.495	2.321
$C1_6H_{34}$	2x 5.419, 2x 4.811	2x 4.270	1.901	2.468	2.082	1.948
$C_{18}H_{38}$	2x 5.414, 2x 4.820	2x 4.285	1.829	2.413	2.018	1.890
$C_{20}H_{42}$	2x 5.425, 2x 4.840	2x 4.293	2.232	2.805	2.535	2.388
$C_{24}H_{50}$	2x 5.384, 2x 4.820	2x 4.289	1.721	2.323	1.928	1.807

Table 2A. Geometric parameters characterizing the triclinic packing chains including the two shortest distances between a given chain and its neighbors (d1 and d2) and the symmetry measures for the hexagon and hexagonal prisms. We use "$\perp$" or N- to indicate that the hexagonal prism has perpendicular chains or has been normalized using the height c respectively.

Chain	d_1(Å)	d_2(Å)	S(L_6-HP-6)	S(HPR-12)	Chain	d_1(Å)	d_2(Å)	S(L_6-HP-6)	S(HPR-12)
$C_{13}H_{28}$	2x 5.104	4x 4.539	0.661	0.619	$C_{35}H_{72}$	2x 4.974	4x 4.484	0.513	0.476
$C_{15}H_{32}$	2x 5.079	4x 4.543	0.594	0.556	$C_{36}H_{74}$	2x 5.104	4x 4.543	0.547	0.508
$C_{17}H_{36}$	2x 4.972	4x 4.483	0.511	0.474	$C_{37}H_{76}$	2x 4.957	4x 4.469	0.511	0.474
$C_{19}H_{40}$	2x 5.038	4x 4.587	0.416	0.393	$C_{38}H_{78}$	2x 5.079	4x 4.543	0.515	0.477
$C_{21}H_{44}$	2x 4.970	4x 4.484	0.502	0.466	$C_{40}H_{82}$	2x 4.972	3x 4.483	0.531	0.492
$C_{23}H_{48}$	2x 4.967	4x 4.473	0.522	0.484	$C_{41}H_{84}$	2x 4.956	4x 4.463	0.523	0.485
$C_{25}H_{52}$	2x 4.961	4x 4.473	0.510	0.472	$C_{44}H_{90}$	2x 5.038	4x 4.587	0.557	0.517
$C_{27}H_{56}$	2x 4.953	4x 4.471	0.498	0.462	$C_{46}H_{94}$	2x 4.967	4x 4.484	0.550	0.513
$C_{29}H_{60}$	2x 4.950	4x 4.468	0.499	0.462	$C_{50}H_{102}$	2x 4.967	4x 4.473	0.520	0.482
$C_{31}H_{64}$	2x 4.934	4x 4.464	0.476	0.441	$C_{60}H_{122}$	2x 4.961	4x 4.473	0.535	0.495
$C_{33}H_{68}$	2x 4.995	4x 4.487	0.548	0.509					

Table 3A. Geometric parameters characterizing the orthorhombic packing chains including the two shortest distances between a given chain and its neighbors (d1 and d2) and the symmetry measures for the regular hexagon and hexagonal prism.

APPENDIX 2: PYTHON PROGRAM

Program to calculate the coordinates of the perpendicular prism according to the following scheme:

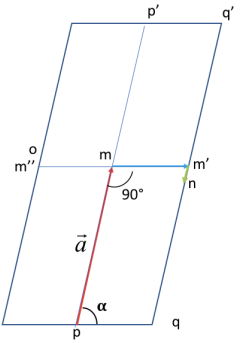

Figure 1A. Illustration of the nomenclature used in the Python program.

```
#-------------------------------------------------------     -----------#
#  This program is used to calculate the coordinates of the perpendicular#
#  hexagonal  prism  from  the  coordinates  of  a  hexagonal  prism  with#
# orientation in its chains TFG Laura Sánchez Muñoz 1/12/2020.          #
#-------------------------------------------------------------------------#
#----------------------------Import modules----------------------------
import numpy as np
import math
import statistics as stats
# --------------------------- Functions -------------------------------
#  Function used to calculate the difference between two points with three
coordinates
def dif(x,y):
```

```
    return [x[0]-y[0], x[1]-y[1], x[2]-y[2]]
#  Function used to calculate the difference vector
def vector(x,y):
    qp = []
    for element in x:
        qp.append([element[0]-y[0], element[1]-y[1], element[2]-y[2]])
    return qp
# Function used to calculate the module of a vector with three cordinates
def modul(x):
    modul_qp = []
    for element in x:
   modul_qp.append(math.sqrt(element[0]**2+element[1]**2+element[2]**2))
    return modul_qp
# Function used to calculate the angle (alpha) between the points m-p-q
def angle(a,qp,ma,mqp):
    alpha = []
    n=len(qp)
    for i in range(n):
       alpha.append(math.acos((a[0]*qp[i][0]+a[1]*qp[i][1]+a[2]*qp[i][2]
)/(ma*mqp[i])))
    return alpha
# Function used to calculate a midpoint
def calculm(x,y):
    m0 = []
    for element in x:
        m0.append([element[0]+y[0], element[1]+y[1], element[2]+y[2]])
    return m0
```

```
# Function used to calculate a vector from trigonometry
def vectorverd(x,y):
    nm = []
    n=len(x)
    for i in range(n):
            nm.append(x[i] *math.sin(math.pi/2 - y[i]))
    return nm
# Function used to calculate the position of a point in space
def ncalcul(x,y,z):
    n = []
    l=len(x)
    for i in range(l):
            n.append(       [x[i][0]-y[i]*z[0],x[i][1]-y[i]*z[1],x[i][2]-
y[i]*z[2]] )
    return n
# Function used to calculate new coordinates
def newcoord(x,y):
    base = []
    for element in x:
        base.append( [element[0]-y[0], element[1]-y[1], element[2]-y[2]]
)
    return base
# ------------------------- Main program -------------------------------
# Read the input file with the coordinates of the hexagonal prism
file = input("What file do you want to read?    ")
output = input("How will you name the exported file?   "
f = open(file, "r")
```

```
nat = int(f.readline())

print("Number of atoms  {:4d}:  ".format(nat))

title = f.readline().strip

# The coordinates are put in a matrix and separated according

# to whether they correspond  to the top or the bottom of the prism

# The bottom is named "q" and the top "q0"

atm_l = []

for atom in range (nat):

    line = f.readline().strip()

    line_l = line.split()

    atm_l.append( [float(line_l[1]), float(line_l[2]), float(line_l[3])])

f.close()

q = atm_l[:6]

q0 = atm_l[6:]

for at in q:

   print(at)

for at in q0:

   print(at)

# Calculation of the central coordinates of the perpendicular prism. Center

coordinates are calculated using the mean of coordinates p and p0

x_p = [atm_l[i][0] for i in range(6)]

y_p = [atm_l[i][1] for i in range(6)]

z_p = [atm_l[i][2] for i in range(6)]

x_p0 = [atm_l[i][0] for i in range(6,12)]

y_p0 = [atm_l[i][1] for i in range(6,12)]

z_p0 = [atm_l[i][2] for i in range(6,12)]

p_center = [stats.mean(x_p), stats.mean(y_p), stats.mean(z_p)]
```

```
print(p_center)

p0_center = [stats.mean(x_p0), stats.mean(y_p0), stats.mean(z_p0)]

print(p0_center)

# Calculation of vector v and a.  Vector v is calculated from the difference between central p and p0

# Vector a is calculated from the value obtained for the vector v multiplied 1/2

v = dif(p0_center, p_center)

v = np.array(v)

a = 1/2 * v

modul_a= math.sqrt(a[0]**2+a[1]**2+a[2]**2)

avectr = a/modul_a

# Calculation of vector Q-P (for all the coordinates q)

# QP is calculated from the difference between q and p

# To get the value of the vector Q-P, the module of the vector is calculated and then the value of alpha is obtained

QP = vector(q, p_center)

print('QP')

print(QP)

modul_qp = modul(QP)

print('modul qp')

print(modul_qp)

alpha_ = angle(a, QP, modul_a, modul_qp)

print()

print('Alpha')

for ang in alpha_:

   print(ang,math.degrees(ang))
```

```
# Calculation of the center midpoint, m, and the midpoint for coordinate
q, m0.
m = p_center + a
m0 = calculm(q, a)
for vv in m0:
   print(vv)
# Calculation of segment m0-m
# The segment m0-m is calculated by the difference between m0 and m
# To get the value of the segment m0-m, the module is calculated
dif_mm = vector(m0 , m)
modul_mm = modul(dif_mm)
print('modul mmp')
print(modul_mm)
# Calculation of green vector. This calculation will be done using
trigonometry
nm0 = vectorverd(modul_mm, alpha_)
print(nm0)
# Calculation of the coordinates of the perpendicular prism
# For the coordinates of the bottom, the value of a is substracted
# For the coordinates of the top, the value of a is added
n0 = ncalcul(m0, nm0, avectr)
print('n0')
for i in n0:
  print(i)
n_q_new = newcoord(n0, a)
n_q0_new = calculm(n0, a)
print(n_q_new)
```

```
print(n_q0_new)

#Calculated data is exported to a .xyz file

document = open(output, 'w')

document.write('{} \n'.format(nat))

document.write('{} \n'.format('title'))

for atom in n_q_new:

    document.write('C  {:10.6f}   {:10.6f}   {:10.6f} \n'.format(atom[0],
atom[1], atom[2]))

for atom in n_q0_new:

    document.write('C  {:10.6f}   {:10.6f}   {:10.6f} \n'.format(atom[0],
atom[1], atom[2]))
```