From dynamics to structure and function of model bio-molecular systems

The research described in this thesis was performed in the section Fundamental Aspects of Materials and Energy of the department Radiation, Radionuclides and Reactors, faculty of Applied Sciences, Delft University of Technology, the Netherlands

From dynamics to structure and function of model bio-molecular systems

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 24 april 2007 om 10.00 uur

door

Fabien FONTAINE-VIVE-CURTAZ

Diplôme d'Etude Approfondies, Chimie-Physique Moléculaire et Structurale,
Grenoble

geboren te Annecy, Frankrijk

Contents

Abbreviations

A: Adenine
AIMD: Ab-Initio Molecular Dynamics
B: Breathing
BZ: Brillouin Zone
C: Cytosine
CI: Configuration Interaction
DFT: Density Functional Theory
DNA: DeoxyriboNucleic Acid
EDTA: EthyleneDiamineTetraacetic Acid
EISF: Elastic Incoherent Scattering Function
FF: Force Fields
FT: Fourier Transform
G: Guanine
GGA: Generalized Gradient Approximation
HB: Hydrogen bonds
HF: Hellman-Feynman
ILL: Institute Laue-Langevin
INS: Inelastic Neutron Scattering
IP: in plane
IR: Infra-Red
IXS: Inelastic X-ray Scattering
LDA: Local Density Approximation
M: Mol/L
MD: Molecular Dynamics
NMA: N-Methyl Acetamide
NMR: Nuclear Magnetic Resonance
NS: Neutron Scattering
OP: out of plane
OS: Out of plane Swinging
PaR: Participating Ratio
PAW: Projector Augmented Wave
PBE: Perdew Burke Ernzerhof
PDA: Pyridine-3,5-Dicarboxylic Acid
PES: Potential Energy Surface
PG: PolyGlycine
PP: PolyProline
PPG_{10}: (Proline-Proline-Glycine) $_{10}$
PPTA: Poly(p-Phenylene TerepthalAmide)
PR: Pure Rotation
PS: Plane Swinging
PW91: Perdew Wang
QENS: Quasi Elastic Neutron Scattering
QM/MM: Quantum Mechanical / Molecular Mechanical
S: out of axis Shifting
SANS: Small Angle Neutron Scattering
SC-loop: Self Consistent loop
SSHB: Short Strong Hydrogen Bond
T: Thymine

TDC: Transition Dipole Coupling
TR: Twisting Rotation
UPA: Urea-Phosphoric Acid
VACF: Velocity AutoCorrelation Function
vDOS: Vibrational Density Of States
VDW: Van der Walls

Chapter 1

Introduction – structure, dynamics and function

The structure and dynamics of every molecular system are determined by the potential energy (hyper-) surface (PES), that is the potential energy variation of the systems as a function of the 3N coordinates, where N is the number of atoms in the system (see Figure 1). Minima of the PES correspond to stable and meta-stable structures and the structure assumed by a system depends on the potential energy of that structure and the number of such structures i.e. entropy.
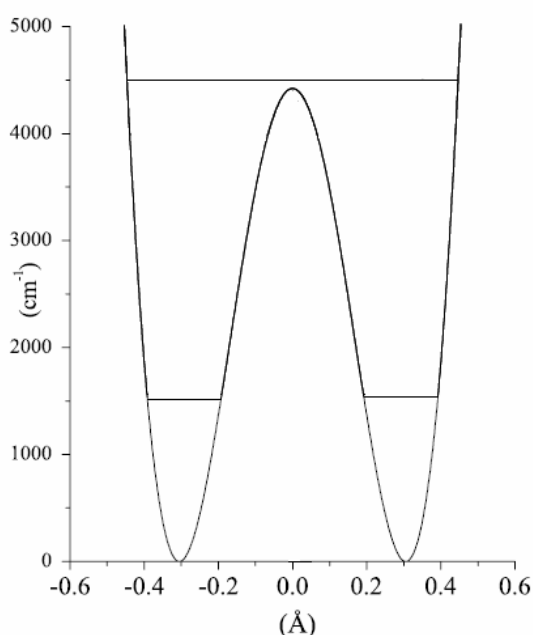
Figure 1

A symmetric two-well potential of hydrogen bond in malonaldehyde [1] illustrating different aspects of a potential energy surface.

Dynamics are determined by the shape of the PES around and between minima. Molecular vibrations are governed by the gradient of the PES in the proximity of the minima, which is often close to harmonic and the corresponding motion is that of a simple harmonic oscillator. A system of N atoms is typically described by a set of 3N normal modes.

Potential energy barriers separate the minima of the PES and energy is required to go over a barrier in a classical diffusive process. Large amplitude motion, as compared to

vibrational amplitudes, which relates the two minima, may correspond to different conformations of a molecule or a chemical reaction (Equation 1) in which, in the simplest case, a proton is transferred along a hydrogen-bond. Such a chemical reaction is an example of the function of a molecular system.
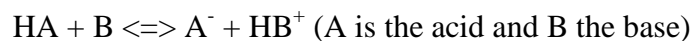
$$HA + B <=> A^- + HB^+ \text{ (A is the acid and B the base)}$$
Equation 1

The case of proton transfer has been extensively studied [2]. If the potential energy of the reactants (A+B) and products (A'+B') is the same, so that the minima of the PES are equally deep (see Figure 1), the proton can pass through the barrier by quantum tunneling and a quantum mechanical wavepacket, localized initially in one well, will oscillate between wells. Normally, however, the reactants and products do not have the same potential energy and energy has to be exchanged with an external source (for example, a heat bath) which results in a dissipative process with a characteristic time rather than a frequency.

The concept of a PES is most useful at low temperature where inter-atomic distances are well-defined and the PES is determined unambiguously. With increasing temperature inter-atomic distances fluctuate and the PES becomes smeared and, in the event of a significant change in coordinates, the PES is considerably modified. For, example, if the PES shown above corresponds to the reactants and products of a proton transfer reaction in an isolated system, then the PES of a neighbouring system will depend, via inter-system interactions, on whether it is surrounded by reactants or products.

Since vibrational dynamics occur in the vicinity of minima of the PES, whereas dissipative dynamics occur in structurally unstable regions of maxima, it is the molecular vibrations which are intimately linked to the molecular structure. Function on the other hand involves a change in the system like the opening of base pairs in DNA in replication or the docking of small, drug molecules in the active sites of proteins and it is the barrier crossing processes that are important.

A range of experimental methods are used to characterize the PES. Diffraction techniques (neutron, X-ray, electron) are used to determine structures corresponding to minima in the PES. Spectroscopy (inelastic neutron and X-ray scattering (INS/IXS), infra-red, Raman …) is used to measure molecular vibrations. Diffusion or reaction rates relating to barrier crossing processes can be measured with techniques like quasi-elastic neutron scattering and NMR.

Whereas diffraction measurements can lead directly to structural details, the determination of normal modes requires a model of inter-atomic interactions since they govern the equations of motion of the atoms via the inter-atomic force constants. The inter-atomic interactions also determine the relative stability of different structures, since they correspond to the residual forces acting on all atoms being zero. Thus it is evident that the PES can be completely determined from a knowledge of inter-atomic interactions.

Experimental data relating to structural stability and molecular vibrations has traditionally been used to refine the parameters in empirical force fields. As computers have become more powerful and the scope of quantum chemistry methods has increased to handle systems of hundreds of atoms, experimental data has provided a reference for different levels of approximation relating to the convergence of these methods. The most computationally efficient implementations of quantum chemistry methods are based on density functional theory (DFT), in which case experimental data is often used, for example, to determine the most appropriate form of the functional that describes the exchange and correlation of electrons.

Experimental data is used to test how accurately the PES can be calculated and once this has been established, vibrational dynamics can be used to select one of a set of potential structures. By calculating the normal modes of vibration for each of the structures, the structure which reproduces the measured spectrum of vibrations is considered to be the relevant structure. If a parameterized force field is used to calculate the PES, the validity of the result may depend on the parameterization and functional forms in the force field. On the other hand, by using where possible a parameter-free, fully converged, quantum chemistry method, the structure-dynamics relation is more reliably established.

In recent years, DFT methods have been shown to be very successful in accurately calculating vibrational spectra for solid state, molecular systems of limited size (≤ 200 atoms) of known structure, including hydrogen-bonded systems [3,4,5,6,7]. By taking the known solid state structure of the molecular system, the set of vibrational modes and, in turn, the spectral profile, are calculated. For this purpose, INS is preferred over IR and Raman techniques since the spectral intensity depends only of the amplitude of the atomic displacements and the corresponding scattering cross-sections. In addition, spectra are rich in peaks since there are no selection rules.

In these cases a reliable assignment of vibrational modes was required. For n-methyl acetamide [3], the hydrogen bond between peptide groups on neighbouring molecules was originally assumed to be "exceptionally" short and accordingly INS data was analysed using a parameterised force field with a force constant giving a low frequency (~1600 cm^{-1}) O-H stretch mode. These calculations were repeated with DFT methods and the hydrogen bond was found to be of normal length with a O-H stretch frequency of ~2500 cm^{-1}, the total spectral profile being well-reproduced. This example illustrates the danger of refining force field parameters.

Another notable example concerns benzoic acid. In this case, the proton transfer reaction rate in dimers had been measured as a function of temperature, revealing an activation energy related to the dimer vibrations which promote proton transfer [8]. These vibrations were identified from DFT-based, normal mode calculations, their displacement vectors being consistent with a theoretical model relating vibrational modes to the reaction coordinate and their frequencies matching closely the measured activation energy [4].

The goal of this thesis is therefore to extend recent work on the structure and dynamics hydrogen-bonded molecular crystals to model systems and processes more closely related to bio-molecular systems. Hydrogen bonds play a fundamental role in many processes in systems of chemical and biological interest. Intermediate in strength between covalent bonds and inter-molecular interactions, they can confer stability on molecular assemblies and yet allow these structures to break-up under specific conditions. DNA is a prime example in which two strands of covalently bonded nucleotides are connected by hydrogen bonding between compatible base

molecules, 3 hydrogen bonds between guanine (G) and cytosine (C) and 2 hydrogen bonds between adenine (A) and thymine (T), giving rise to the famous double helix structure [9]. The stability of DNA under different physical conditions (temperature, humidity, pH …) is vital to the reliability of the genetic coding contained in DNA. And yet the strands of DNA also have to split in different biophysical processes, for example during replication in which the genetic code is copied into pairs of DNA molecules.

At the end of this thesis (chapter 4, "DNA"), we present an as yet unpublished investigation of the molecular vibrations of DNA. While DFT calculations are now tractable on hydrated structures of DNA [10,11], which contain more than 1000 atoms, the number of such calculations that would have to be performed in a normal mode calculation is prohibitive. We have therefore used the CHARMM force field and our results are intimately related to the parameters of the force field. These calculations do however correctly reproduce the measured dispersion of low frequency vibration modes and give insight into vibrational modes related to base-pair opening, a key reaction coordinate, as in benzoic acid. To do this, we have rewritten the phonon code [69] used in previous work to handle an unlimited number of atoms.

The strength of a hydrogen bond is variable depending on electrostatic, dipolar, dispersive, and in some cases, covalent interactions. As a result hydrogen bonds are broadly classified as varying from weak to strong, the latter having the shortest donor-acceptor distances (~2.4 A) and protons that are almost centred between heavy atoms. Short, strong hydrogen bonds (SSHBs) cannot be accessed directly in complex molecular systems since the resolution in diffraction experiments is not sufficient to distinguish the small geometrical differences between weak, normal and strong bonds and, computationally, precision also has to be sacrificed in order to deal with a large number of atoms. Since SSHBs are thought to mediate important biophysical processes such as stabilizing the intermediate state of certain enzymatic reactions [12,13], understanding their physical properties is of considerable importance. We have focused on a series of model systems with SSHBs that exhibit complete or partial transfer of the proton across the hydrogen bond as a function of temperature. Since the donor and acceptor molecules change chemical composition upon proton transfer, this is an example of a simple chemical reaction, which can be regarded as the

function of these molecular systems. In this way we have studied the structure-dynamics-function relation in SSHBs crystals. In order to study temperature dependent structural changes we have had to extend our use of ab-initio calculations to include molecular dynamics (AIMD) simulations in which temperature describes the average kinetic energy of the atoms of the system. This published work is presented in chapter 2 ("Short strong hydrogen bonds") as a summary of 2 papers, which are included in full in annexes B and C.

Another significant extension undertaken here to previous work in vibrational spectroscopy of molecular systems studied by INS and DFT is from molecules to polymers. This is an important step towards dealing with bio-molecular systems since they are generally polymeric, for example, proteins and DNA. In addition hydrogen bonding plays a key role in bio-molecular structure, stabilizing secondary structures such as beta sheets and helices [14]. It is in this context that the structure-dynamics relation is widely exploited in order to obtain structural information, mainly from IR spectra. In complex biomolecular systems, detailed structural information may be lacking due to difficulties in crystallization or simply due to the size of the molecular entities. In these cases, vibrational spectra and, in particular, the profiles of the so–called amide bands, which concern mainly the vibrations of the amide group C(=O)-N-H, are used to determine and quantify the presence of secondary structures. The basis of this method of analysis is the direct link between hydrogen bonding, amide groups and secondary structure.

We have investigated this structure-dynamics relation in a model compound, PPTA (known commercially as Kevlar), in which hydrogen bonding gives rise to the formation of 2-dimensional sheets, that are similar to beta sheets. In Kevlar, despite the obvious interest in this compound, structural questions persisted until recently concerning the packing of the 2-D sheets in the solid state. Neutron fibre diffraction on partially deuterated samples gave clear information on the overall, relative alignment of the 2-D sheets. We used INS and, in particular, the amide-V vibrations, to gain insight into the relative orientation of amide groups between the sheets, a structural issue that was unclear from diffraction work. These results have been published (annexe D) and are presented briefly at the start of chapter 3 ("Biopolymers").

Having explored the structure and vibrations of a number of hydrogen-bonded molecular systems in the solid state and observed the subtle dependence of vibrations on structural details, the analysis of amide vibrations in complex bio-molecular systems seemed to be worthy of further investigation with modern DFT methods. INS data exists for polyglycine (PG), which can crystallize in either a beta-sheet structure (PG-I) or a helical structure (PG-II), and polyproline, which also crystallizes in a helical structure (PP-II). These systems and Kevlar provide the basis for a detailed study of the variation of amide vibrations for sheet and helical, secondary structures in different solid state environments. This unpublished work is presented in chapter 3.

The final biopolymer to be presented in the scope of this thesis is collagen, which is the main protein found in connective tissue, such as cartilage, ligaments, tendons, bone and teeth, in mammals. Collagen is therefore not a model system. It presents not only a helical secondary structure, but also a tertiary structure in which three helices are intertwined. INS spectra were measured by Middendorf and co-workers and analysed without recourse to molecular modelling [39]. A number of reasonable conclusions were made concerning vibrational assignments but the assignment of modes characteristic of the triple helix is certainly more speculative. Using normal mode and molecular dynamics simulations, we have looked in detail at the vibrations of collagen.

Collagen also presents another key feature of bio-molecular systems, namely the protein is surrounded by water, which has well-defined crystalline positions between and close to helices, forming a first hydration layer, and then bulk water further away from the protein. Water is included in all DFT calculations. Amorphous water in a numerical DFT calculation results in a poorly defined global minimum and, as a result, a significant number of the lowest frequency modes in a normal mode calculation are negative. This problem is overcome by using a force field, in which the total energy of the system is given by an analytical expression, affording much higher numerical precision. Of course, the force field (CHARMM as used for DNA) introduces an uncertainty in all results related to its parameters. This unpublished work concludes chapter 3.

This thesis work is both experimental and numerical. New INS data has been measured on spectrometers at ILL and ISIS for SSHB systems, Kevlar and DNA. In contrast to previous measurements performed on powders, Kevlar and DNA were used in the form of aligned fibres, which allows for a higher level of experimental information to be obtained. While Kevlar fibres can be purchased, DNA fibres have to be spun and for this purpose a spinning machine was transferred from Juelich to ILL and modernized as part of this thesis project. Numerically, we have extended the use of DFT methods to polymers and worked on bigger systems, up to 350 atoms for collagen. In the case of DNA where the large number of atoms means that DFT cannot, in practice, be used, we have used the CHARMM force field. In order to perform large scale phonon calculations, the code used to data (PHONON), has been rewritten in collaboration with Dr Franci MERZEL (Ljubljana, SLOVENIA). A key aspect of the use of DFT and force field methods in this thesis is the critical appraisal of their strengths and limitations. INS and numerical simulations, especially using DFT, are the key techniques used in this thesis. Readers not familiar with these techniques should consult A.

The thesis is therefore presented in the form of this introduction and three chapters; short hydrogen bonds, biopolymers and DNA. Where work has already been published, the contributions to these chapters are concise and the full articles can be found in the corresponding annexes (B - D). Other annexes contain details of the numerical methods used in this work (A) and details of the spinning machine that was installed at ILL as part of this thesis in order to produce aligned DNA samples (E).

Chapter 2

Short strong hydrogen bonds crystals (see Annexes B-C)

II/1/ Introduction

SSHBs are thought to play a key role in some bio-physical processes in stabilizing the intermediate state of certain enzymatic reactions through a subtle modification of the local environment of the hydrogen bond [12]. In the enzymatic reaction scheme of the ketosteroid isomerase, the transition state is stabilised by the formation of SSHB to allow an allylic rearrangement in the substrate involving a double intramolecular proton transfer [13]. However, studying SSHBs in detail in complex systems is difficult, experimentally due to the limited resolution of diffraction techniques and, numerically, due to the precision of methods that has to be reduced in order to handle a large number of atoms. Model systems are therefore required. Neutron diffraction has been used to study a number of systems with SSHB that show proton transfer as a function of temperature, up to room temperature. In the context of this thesis, we regard these systems as an example of the structure-dynamics-function relation. Proton transfer is the function, which results from structural modification induced by temperature. Details of the mechanism are not clear and there are several possibilities.

1. Thermal fluctuations cause lattice expansion which may modify the geometry of the SSHB in a way that favours proton transfer.
2. The potential well may contain a low lying, excited vibrational level, which could be populated at room temperature.
3. The high temperature structure is entropically stabilised by lattice phonons, which results in proton transfer.

Previous INS and DFT work on hydrogen bonded systems, shows how these questions can be addressed directly by simulations. However, since SSHBs have not been widely treated by DFT methods, it is essential to have INS data to help identify the critical vibrations involving the hydrogen bond and, more generally, to check the accuracy of the DFT calculations. Furthermore we have to go beyond the harmonic approximation, which means mapping out the PES and analyzing the vibrations in potential energy wells by solving the corresponding Schroedinger equation. Molecular dynamics simulations are particularly useful in this respect since the harmonic

approximation is removed and temperature is introduced via the kinetic energy of the atoms.

The work on four particular SSHB systems presents some common ground with recent work in terms of applying INS and DFT to study hydrogen-bonded molecular crystals, but also includes new technical aspects like the use of MD in parallel with normal mode analysis and the exhaustive mapping of the PES and analysis of anharmonicity. The model systems presented here are the crystals of pyridine-3,5-dicarboxylic acid (PDA) and urea-phosphoric acid (UPA), which have revealed the most interesting results about the mechanism of proton transfer as well as perspectives for further neutron experiments.

II/2/ Structural information

In these systems the SSHB are of the type O-H…O for UPA, compared to PDA which has a N-H…O SSHB (Figure 2). Thus, proton transfer is not chemically limited to a particular type of hydrogen bond.
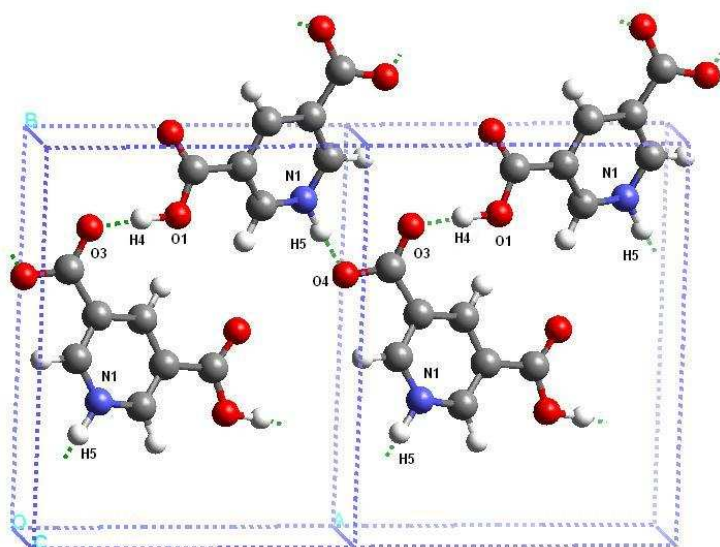
Figure 2

An overall view of the low temperature crystal structure of PDA as reported in [15,16].

Neutron diffraction provides the reference crystallographic data, which is the basis of all simulations presented here. Solid state DFT methods tend, in geometry optimisation and low temperature MD calculations, to overestimate the donor-

acceptor distance and underestimate the donor-H distance, which hinders the proton migration effect in the simulations (Table 1). In PDA, thermal fluctuations cause lattice expansion along the c axis and thus an increase of the inter-sheet distance which does not significantly affect the SSHB geometry. Geometry optimisation in any measured unit cell (low or high temperature) results in a lowest energy structure with a short N-H distance (~ 1.1 Å). Thermally induced lattice expansion is therefore not the mechanism driving proton transfer.

	N1….O4	N1….H5	H5….O4	N1-H5-O4
exp data				
296K	2.525	1.308	1.218	177.9
100K	2.526	1.181	1.361	172.3
15K	2.523	1.213	1.311	176.4
MD simulations				
300K	2.541	1.206	1.331	169.3
250K	2.546	1.162	1.371	170.1
15K	2.535	1.128	1.409	174.2
optimised geometries				
15K	2.551	1.122	1.429	176.5

Table 1

Proton transfer as a function of temperature in the N1...H5...O4 SSHB in PDA and obtained from simulation.

II/3/ Vibrational properties

In these systems, the strength characteristic of the HB type is clearly evaluated by calculating the vDOS of the hydrogen atom in short, strong or in normal HBs. The similar results obtained from MD and normal mode calculations in Figure 3 allow us to conclude that anharmonic corrections due to the potential energy variation along the normal coordinates (Figure 4) are small. Furthermore, Figure 3 demonstrates that the stronger the HB, the smaller the spectral range of the hydrogen atom vibrations.
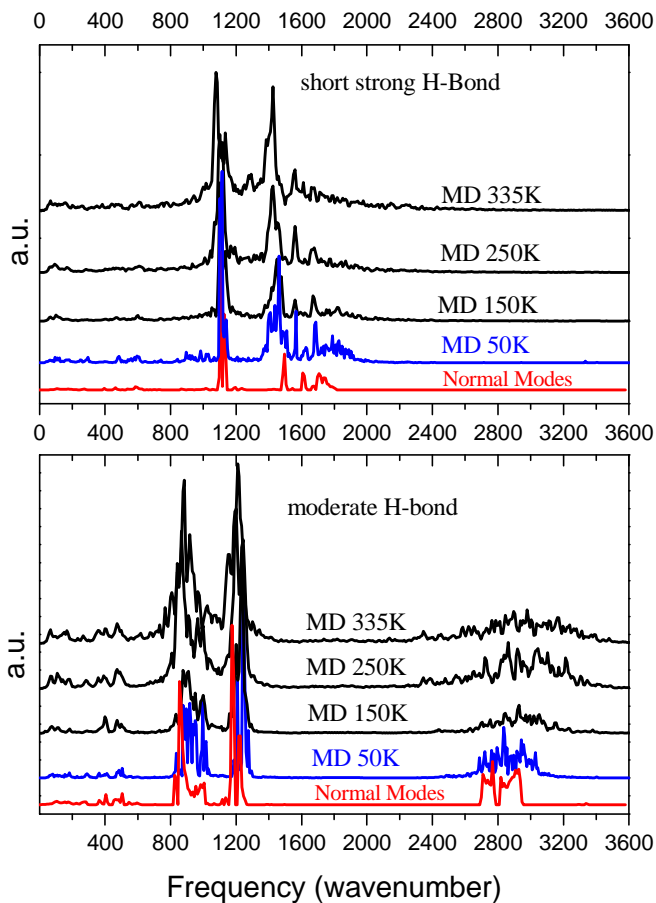
Figure 3

The vibrational density of states of the hydrogen involved in the strong H bond O5-H4...O4 (upper panel) and of the H involved in a moderate H bond O1-H1...O2 (lower panel) in UPA. The lowest curve in each panel is derived from phonon calculations, the upper four curves are from MD simulations.
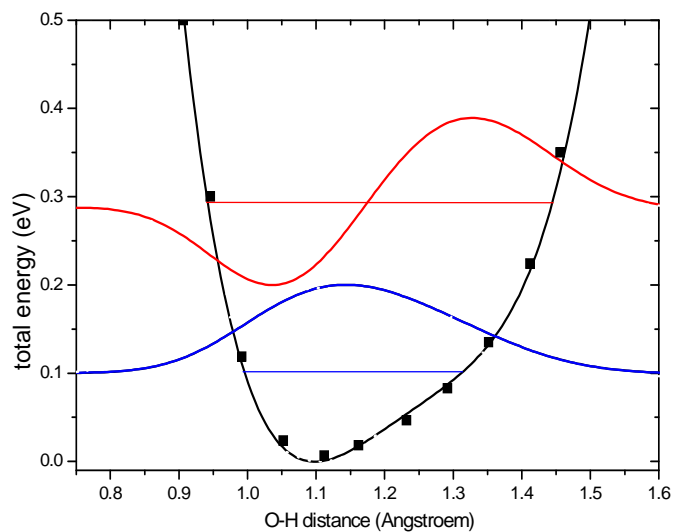
Figure 4

Potential energy well, vibrational energy levels and wavefunctions of the hydrogen atom in the SSHB. The wavefunctions represent the fundamental state n=0 (lower curve) and the first excited state n=1 (upper curve) in the anharmonic potential energy well, the vibrational energy is 1525 cm^{-1}.

Figure 4 shows the potential energy variation along the O-H direction of the SSHB in UPA. The potential well is clearly anharmonic, but the coordinate is not a normal coordinate, the low frequency vibrations of SSHB allowing the O-H stretch to mix with other vibrational modes. Solving Schroedinger's equation for this potential gives an excited state wavefunction with a maximum probability shifted towards to the centre of the HB, but the energy of this excited state means that it cannot be significantly populated at room temperature. Proton transfer is not, therefore, driven by excited vibrational states.

II/4/ Phonon driven proton transfer

One significant difference between the N-H…O SSHBs in PDA and the O-H…O SSHB in UPA is that, experimentally, the protons in PDA migrate beyond the centre of the hydrogen bond. For PDA, the time dependence of the N-H distance shows almost periodic hopping of the proton from one side of the hydrogen bond to the other and the histogram clearly shows two peaks separated by ~0.1 Å, which corresponds to the proton transfer distance measured in the experiment (Figure 5).
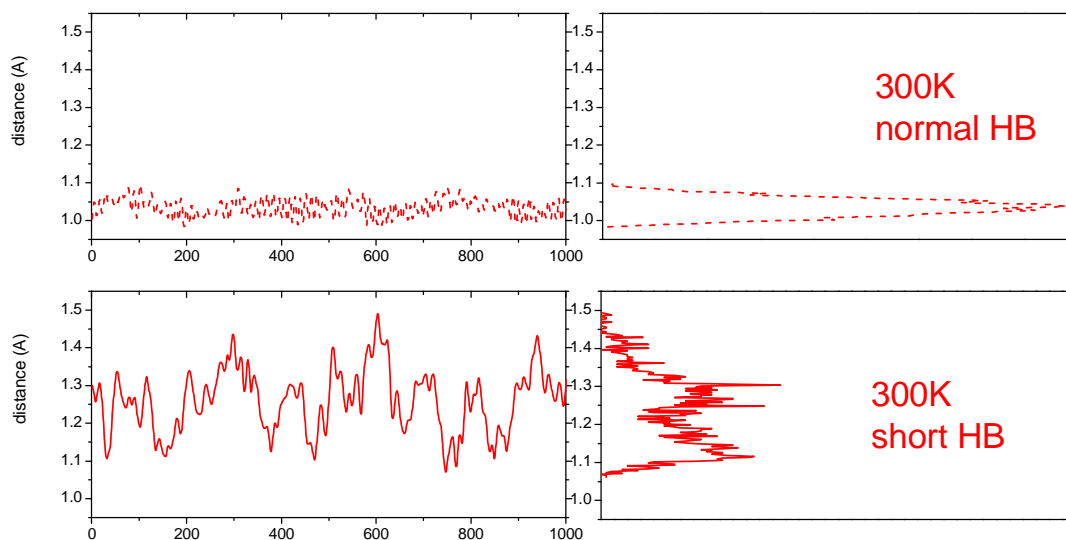
Figure 5

N-H distance (lower two panels) and O-H distance (upper panel) as a function of time over 1 ps. The histograms (right) are integrated over the whole MD trajectories.

Analysing structures for PDA from the MD trajectories show that the lattice vibrations tend to stabilise the proton position on either side of the SSHB, but these positions are not positions of minimum potential energy. In the case of UPA the centred position is stabilised in the MD simulations, corresponding to a partially transferred proton, as measured. The periodicity is around 100 – 200 fs and the Fourier transform (Figure 6) shows an initial peak at low frequency (< 50 cm^{-1}) and two well-defined peaks at 115 and 145 cm^{-1}.
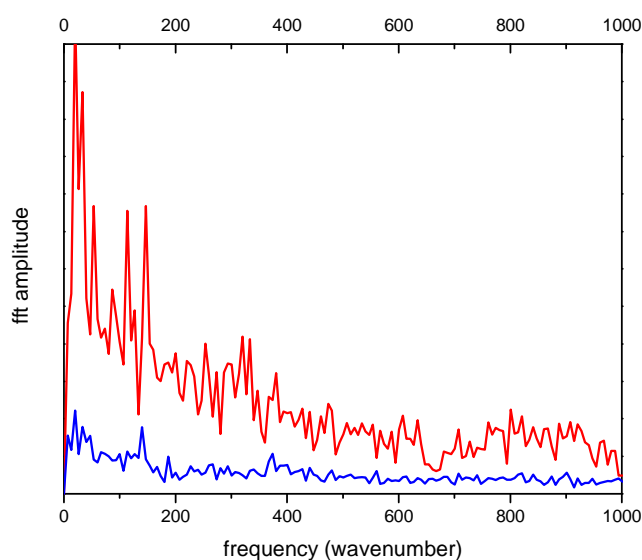
Figure 6

The Fourier Transform amplitude of the time-dependent N1-H5 bond length variation at 300K (upper curve), and 15K (lower curve).

In Figure 7, INS and phonon calculations for PDA are seen to compare well. Analysing the low frequency vibrations, and allowing for anharmonicity, show that the prominent frequencies of 115 and 145 cm^{-1} from figure 5 correspond to vibrations in the plane of the molecular sheets which give rise to the strongest modulation of the hydrogen bond geometry (see Figure 8). These bi-dimensional specific modes are a direct consequence of the molecular sheets in PDA.
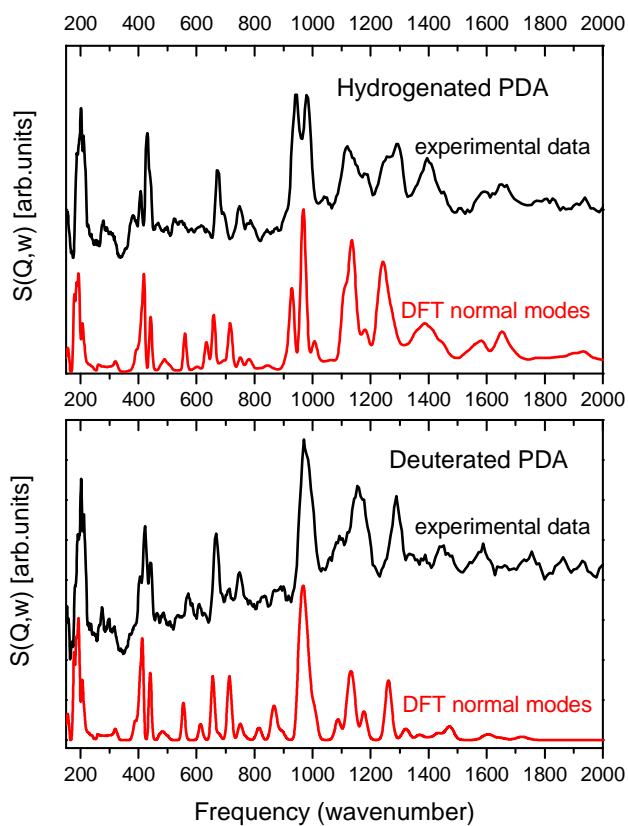
Figure 7

Comparison of measured (upper curve) and calculated (lower curve, normal modes analysis) INS S(Q,w) spectra of fully protonated (upper panel) and partially deuterated (lower panel)pyridine-3,5-dicarboxylic acid (PDA).
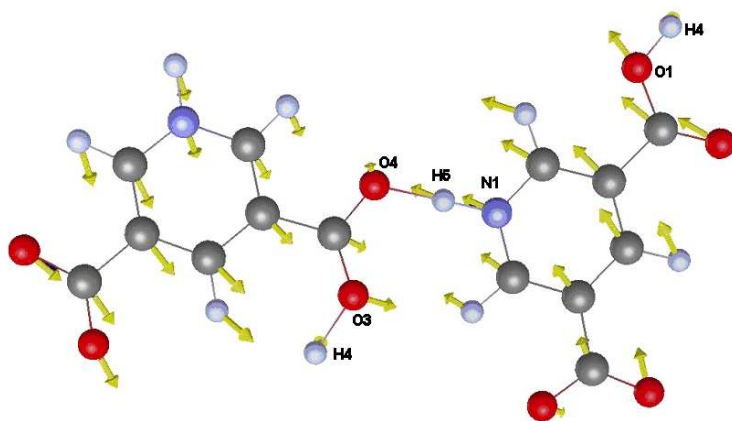
Figure 8

Representation of the vibrational mode of hydrogenated PDA at 155 cm^{-1}.

II/5/ Conclusion

The four systems studied in two published papers offer considerable insight into the function of temperature-driven, proton transfer in SSHBs. In 3 out of 4 cases, the

15

observed proton transfer is reproduced (semi-) quantitatively, but in the fourth sample, the optimized, equilibrium length of the hydrogen bond is too long, that is we no longer have a SSHB. MD has to be used to investigate temperature-induced, structural fluctuations and this method is validated by comparing the vibrational density-of-states from MD, with that obtained from the tried-and-tested normal mode approach.

We have shown that low frequency lattice vibrations stabilize an average crystal structure in which the proton is transferred, towards or beyond the centre of the hydrogen bond. By Fourier transforming geometrical quantities extracted from the MD trajectory, characteristic frequencies for the modulation of the hydrogen bond structure are revealed in the simplest system, PDA. By combining MD and normal mode analysis, the relevant vibrational modes have been identified. In the large number of modes in an organic crystal, one would expect to find vibrational modes which modulate the SSHB geometry and it remains to be investigated why most SSHBs do not show proton transfer.

Finally the case of PDA is particularly interesting since the proton is completely transferred from one molecule to the other. The MD simulations show clearly a two-site jump across the centre of the hydrogen bond, the central position being less stable. In the reference frame of the unit cell and, therefore, in the diffraction experiment, this splitting of the proton density is not observed. This aspect of the MD result could however be investigated experimentally using QENS or by measuring the momentum distribution using Compton scattering techniques [17].

Chapter 3

Structure and dynamics of biopolymers

III/1/ Introduction

Technically, chapter 2 showed advances in the use of DFT methods in the analysis of INS data, but the samples were molecular crystals as studied in earlier work. In this chapter, we apply our panoply of methods to study polymeric systems, including water, which is an important step towards treating real biopolymers with INS and DFT. Indeed, this chapter concludes with a detailed analysis of hydrated collagen which is the main protein found in connective tissue.

Low frequency vibrations of biopolymers and their softening with increasing temperature is a major factor in biological activity. Theoretical work has shown that picosecond fluctuations in the secondary structure are crucial in determining reactions rates in proteins. For example, it has been calculated [18] that ligand binding to an enzyme causes softening of low frequency, delocalized modes. Moreover, Zhou et al. [19] have shown that structural fluctuations, due to collective modes around the active site on a picosecond timescale, is responsible for an enzyme to achieve its substrate specificity. It is well known that biological function of a protein is coupled with anharmonic protein dynamics. Frauenfelder et al. [20,21] have shown that biological function of proteins can be explained by switching motions of molecular groups in response to external factors such as ligand binding. These structural transitions between the resting and the activated states are the consequence of the highly anharmonic dynamics on the rugged potential energy surface of the protein molecule. Other authors [22,23] have shown how, due to the photodissociation of the ligand in myoglobin, vibrational energy is transferred via coupling with the anharmonic protein dynamics. This transfer is achieved by two distinct mechanisms depending on temperature which therefore constitutes a predominant factor for the energy transport in biological systems.

In protein science, hydration is also a very important factor for stability and function [24]. A number of previous studies [25,26] strongly suggest that the dynamical transition of proteins, which is the transition from a low temperature harmonic regime to an

anharmonic one at higher temperature, is intimately linked to the solvent dynamics. This transition is a direct consequence of the large-amplitude anharmonic atomic motions of the protein and these motions, necessary for biological activity, are very dependent on the degree of plasticizing, which is determined by the level of hydration [27]. Monte Carlo analysis of SANS data of hydrated proteins has shown a 10% increase in the density of the first hydration shell compared to bulk water [28]. In terms of dynamics, MD simulations have also demonstrated that the translational motion of water molecules is crucial in allowing large amplitude structural rearrangement [29]. By measuring different concentrations of penta-alanine oligomer's solutions [30], ranging from dry powders to fully hydrated systems, under ambient conditions, the dynamics of the first hydration shell have been shown to be limited to slow rotational reorientation while at higher levels of hydration, translational dynamics dominate, ressembling bulk water.

High frequency vibrations of biopolymers are generally not involved in biophysical processes. They are however exploited in the context of the structure-dynamics relation. For weakly charged systems, molecular vibrations depend on local interactions and are therefore a probe of local structure. This is particularly useful when traditional diffraction methods are not applicable due to the lack of long range order, which is often the case for biopolymers, particularly in solution. Thus IR spectra, especially the amide-I band, are commonly used to quantify the presence of beta sheets and helices. The use of vibrational spectra to identify different polypeptide chain conformations dates from the work of Elliott and Ambrose [31], who observed that the frequency of the infrared-active Amide I vibration of α polypeptides was about 20 cm^{-1} higher than that of β polypeptides. Further studies on synthetic polypeptides [32] established this observation as a firm empirical rule which can be applied for determining the presence of secondary structures in proteins [33, 34]. Topical applications of this method include the study of prion diseases [35].

The major factors responsible for conformational sensitivity of the amide bands include hydrogen bonding and coupling between transition dipoles. In particular, the transition dipole coupling (TDC) leads to the splitting of the amide I mode and the magnitude depends on the orientation and distance of interacting dipoles and thus provides information about geometrical arrangements of the peptides groups [36].

Experimentally, this relationship between the position of the amide band and the type of secondary structure is very useful for homopolypeptides that fold into well defined and homogeneous (i.e. purely α-helical or purely β-sheet) secondary structures [37]. However, proteins fold into a complex three dimensional arrangement which consist of a variety of domains containing polypeptide segments in different types of secondary structures. Since all these conformations contribute in the frequency range of the amide I band of the protein from 1600 to 1700 cm^{-1}, the observed shape and intensity of this band consists of many overlapping bands related to the different structural elements. This give rise to fundamental difficulties since the width of these bands is greater than the separation between the maxima of adjacent peaks. Thus the interpretation requires an extensive mathematical treatment of the experimental data as seen in the IR spectra of the ribonuclease (Figure 9 and Table 2).
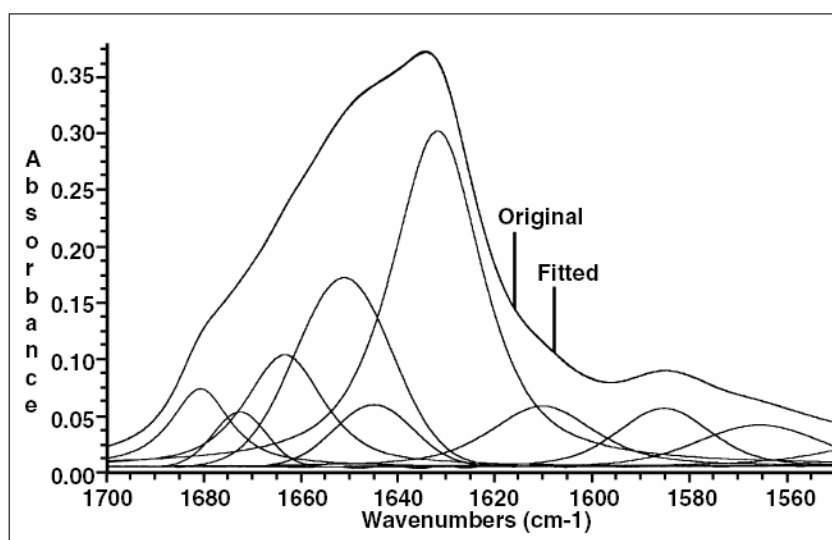
Figure 9

IR spectra (original and fitted) of ribonuclease. The Lorentzian and Gaussian are produced from a complex curve fitting process [38].

No. of Peaks	Center (cm^{-1})	Height	Width	% Lorz	Area	% Area	Assignment
1	1681	0.069	14.7	0.881	1.351	7.3	Turn
2	1673	0.049	12.9	0	0.671	3.6	Turn
3	1663	0.099	19.1	0.658	2.430	13.2	Turn
4	1651	0.167	23.5	0	4.146	22.7	α-helix
5	1645	0.055	19.2	0	1.127	6.1	Random
6	1632	0.297	22.1	0.703	8.642	47.0	β-sheet

Table 2

The peak parameters, areas and assignments of individual component bands in amide-I band of ribonuclease A.

Krimm [37] has demonstrated that normal mode calculations of the amide vibrations is a very powerful method to identify the three dimensional arrangement of polypeptide chains. However, the parameterization of empirical force fields to calculate inter-atomic interactions can be complex. In particular, the magnitude of the splitting is only reproduced by taking into account the transition dipole coupling (TDC) and parameter-free DFT based calculations are, therefore, very useful in this context.

DFT and INS work on molecular crystals reveals a subtle dependence of molecular vibrations of local structure and, furthermore, in large systems, vibrational modes are very complex. The extent to which the structure-dynamics relation is exploited in bio-molecular systems seems, at first sight, to be surprising and, in this chapter, we explore this relation for PPTA (Kevlar) and some simple polypeptides (polyglycine and polyproline). Glycine and proline are the principal components of collagen and, in this case, we look, in particular, for the signature of triple helix formation [39]. In addition, for collagen, the triple helix is hydrated and water is an important component of our molecular models. The low frequency dynamics and the effect of temperature and hydration will be studied in detail.

III/2/ Secondary structures of proteins

III/2/1/ Kevlar, a model for sheet packing in the solid state (see Annexe D)

Poly(p-phenylene terepthalamide) (PPTA) is a polymer which crystallises in the form of beta sheets, peptide linkages in the polymer allowing hydrogen bonding between polymer chains to form 2D sheets (Figure 10). PPTA (commercially available as Kevlar®) is a fiber with exceptional mechanical properties that are due in part to a combination of crystalline and amorphous regions in the sample.
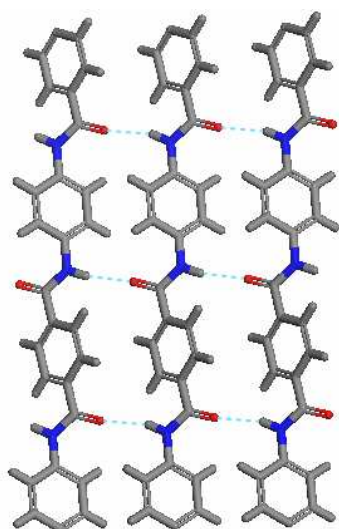
Figure 10

Beta-sheets structure common to all PPTA structures.

The fibre used in this work has a crystallinity of ~70%, making vibrational spectroscopy an important tool for structure investigation. Indeed, uncertainty persists concerning the packing of sheets in the crystal structure of PPTA, which is governed by weak intermolecular interactions. Several structures presenting identical sheet structures but different packing of phenyl and amide groups will be considered (Figure 11). The Northolt [40,41] and Pb [42] structures have an approximately parallel alignment of amide linkages; the Northolt-translated [43] and Liu structures [44] have an antiparallel alignment. The packing of phenyl rings between sheets in all structures is of herringbone type, except in the Northolt-translated structure in which the phenyl groups are stacked with a similar, parallel orientation.
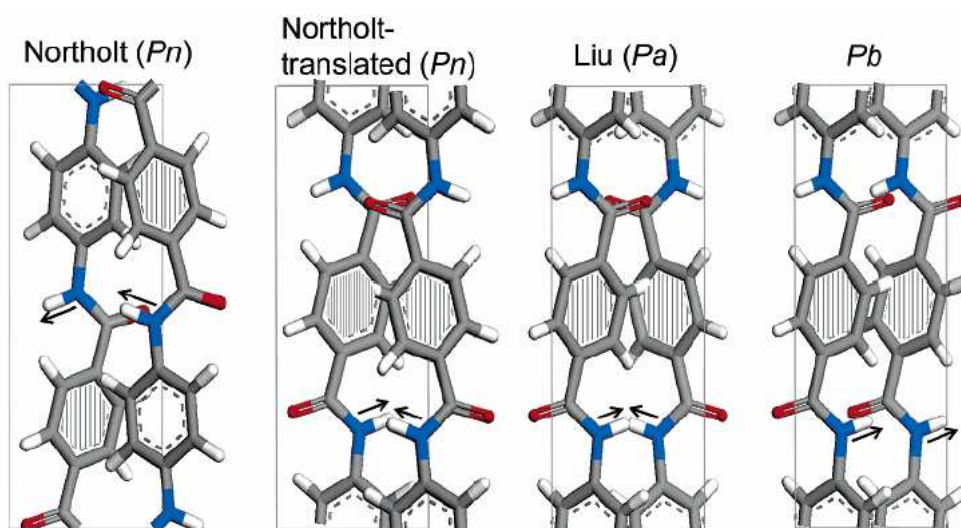
Figure 11

The principle PPTA structures studied in this paper. The two chains in the unit cell are shown, that is, one from each of the neighbouring sheets. Terephtaloyl residues are shaded to show the alignment of the chains. Arrows on the N-H bonds indicate the orientations of the amide linkages.

It is the accurate DFT-based calculation of the weak inter-sheet interactions, in the presence of stronger intra-sheet interactions that enables the structures to be distinguished. By comparing the total energies of the optimized structures, calculations clearly exclude the parallel packing of phenyl groups in the solid state. However, identifying the relative orientation of amide linkages is more subtle because the packing of the amide groups gives much smaller energy differences. Vibrational spectroscopy, coupled with DFT calculations, is therefore an important tool in clarifying the structural issue of PPTA. Spectra have been measured using INS and calculated for four distinct crystal structures (Figure 12). The calculation for the Liu structure is in best agreement with the measured spectrum. The pronounced triplet of peaks between 600 and 1000 cm^{-1} is characteristic of an anti-parallel alignment of amide linkages, which is common to the Northolt-translated structure. The N-H wagging mode, characteristic of the Amide V mode, is the principle probe of the inter-sheet packing because the hydrogen atom can have a large amplitude motion (when this mode is pure), giving an intense peak for parallel aligned amide linkages (Figure 12). Moreover, Kevlar fibres were used in this work and Figure 13 shows that, for the Liu structure, the measured and calculated variations in the INS intensity with fibre orientation match well.
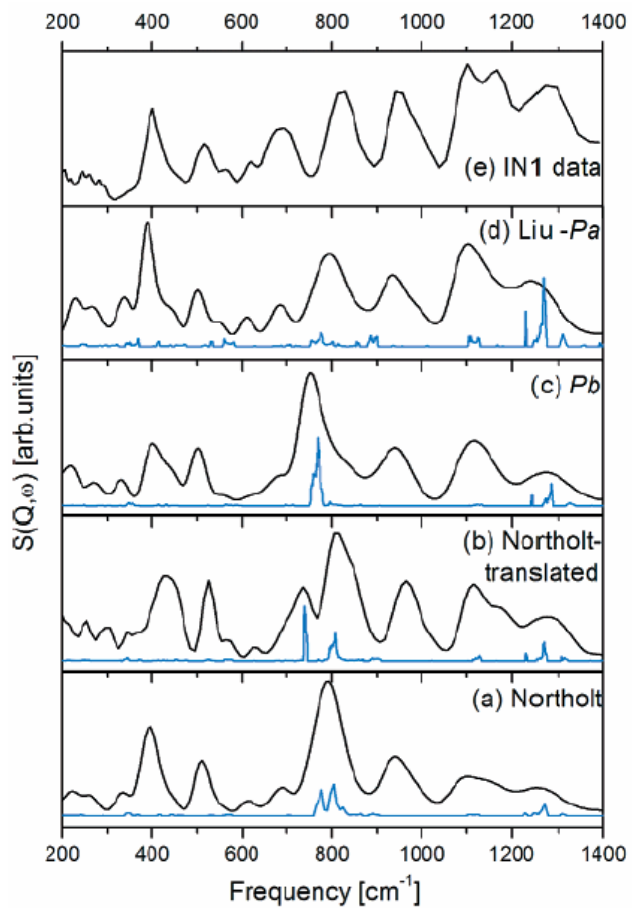
Figure 12

The measured INS spectrum of PPTA, summed over fibre orientations ranging from parallel to the momentum transfer Q vector to perpendicular (top, e). Calculated INS spectra for the same, average fibre and instrument configuration are shown in the lower four panels. The unconvoluted, partial scattering function for the amide hydrogen is also shown in each panel.
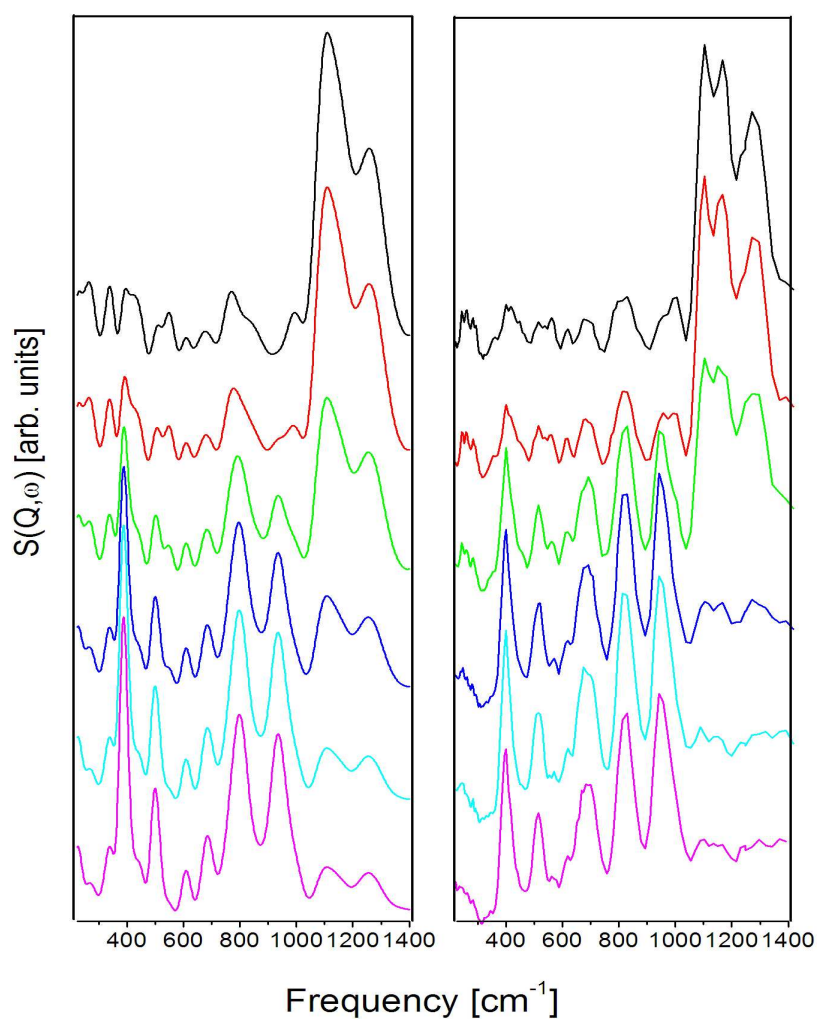
Figure 13

Comparison of calculated (left) and measured (right) and INS vibrational spectra of PPTA for fibre orientations varying from parallel to the momentum transfer vector Q (top) to perpendicular (bottom)

III/2/2/ Polyglycine-I, II and polyproline-II: beta-sheet vs. helices dynamics

The aim of this section is to extend our INS and DFT calculations and the structure dynamics relation to simple helical polypeptide models. The validation of these methods will allow us to go further in the study of more complex polymeric systems with triple helices.

In the solid state, the amino acid chain of polyglycine crystallises in the beta-sheets (PG-I) [45] or alpha-helices (PG-II) [46] structures (Figure 14) and constitutes therefore

a system of choice to evaluate vibrational differences induced by the three dimensional form of the secondary structure. The vibrations of the amide linkages are of particular interest and they are classified as follows.

I: C=O stretch+N-H in-plane bend

II: N-H in-plane bend+C-N stretch

III: C-N stretch+N-H in-plane bend

IV: C=O in-plane bend+ C-C stretch

V: N-H out-of-plane bend+C-N torsion

VI: C=O out-of-plane bend+N-H out-of-plane bend
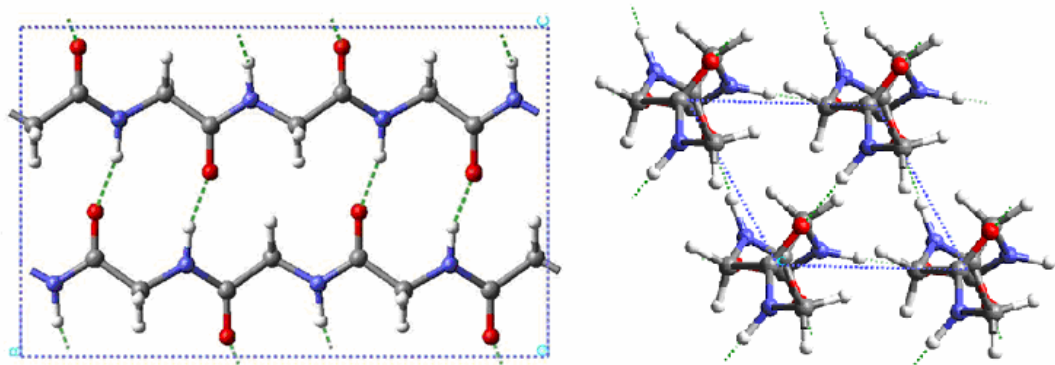
Figure 14

Left: sheet structure of polyglycine I, right: helical structure of polyglycine II

The structure-dynamics relation can be established by determining the vibrational signature of the amide I vibrations (CO stretching and NH in pane bending) because this particular mode presents spectral differences depending on the secondary structure. IR measurements have shown that the amide I band is at 1633 and 1688 cm^{-1} for PGI and 1646 cm^{-1} for PG-II [47]. Compared to experimental data, the well reproduced INS spectra with DFT-normal mode calculations (Figure 15 and Figure 16) allow a validation of numerical models, force constants, and assignment of the vibrational modes. It is shown that without any refinement of force constants, the magnitude of the splitting depending on the secondary structure as well as the position of the amide I band are quite well reproduced by DFT-based calculations (Table 3).

		Amide I	Amide II	Amide III	Amide V
PPTA	exp	1665	1555	1270	720
	calc	1623-1636	1525-1526	1260-1273	729-730
PP II	exp	1657	1415-1460	1238-1246	745
	calc	1634-1648	1415-1442	1289-1298	741-743
PG I	exp	1633-1688	1520	1236	708
	calc	1620-1714	1535-1549	1247-1274	724-732
PG II	exp	1646	1550	1247-1282	707
	calc	1624-1631	1545-1548	1237-1280	702-709; 755-758
PPG$_{10}$	exp	1650	1550	1250	550-590
	calc	1640-1705	1524-1551	1240-1260	575-775

Table 3

Comparison of measured (see reference [48] for PPTA) and calculated amide bands for PPG$_{10}$ and model polypeptides. The description of amide bands in term of vibrational modes are described below.
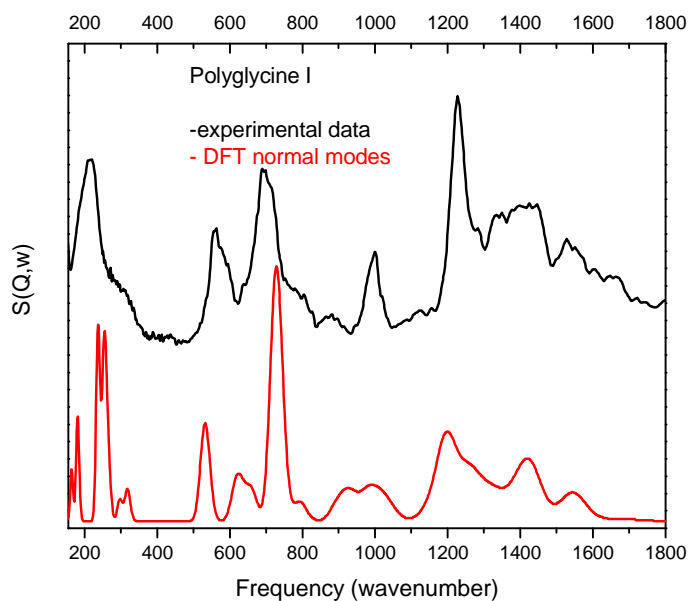
Figure 15

The measured (top) and calculated (bottom) INS spectra of Polyglycine I.
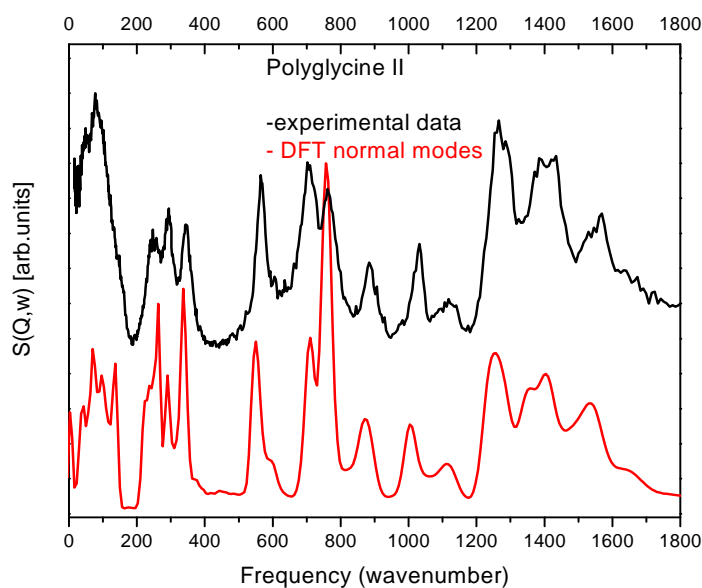
Figure 16

The measured (top) and calculated (bottom) INS spectra of Polyglycine II

Proline and glycine are the amino acids constituting the collagen protein. The natural investigation of the relation structure-dynamics is therefore to extend our methods to the helical structure of polyproline II (PP-II) determined by Arnott and Dover [49]. Information about the amide bands, suggested by experiments [50] and by DFT calculations are presented in Table 3. It is important to notice that proline is the only amino-acid which does not have an amide hydrogen, the nitrogen being included in the ring. It means that the secondary structure is not stabilized by hydrogen bonds between amide groups. Nevertheless, comparison between experimental and calculated INS spectra in Figure 17 allows us to attribute the amide bands.
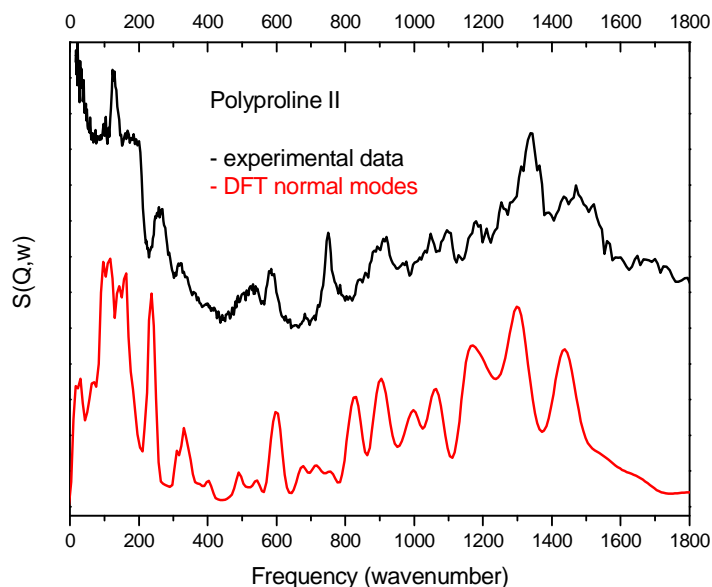
Figure 17

Comparison of measured (top) and calculated (bottom) INS spectra of polyproline II.

This work on PG-I, PG-II and PP-II demonstrates that normal mode analysis, with force constants determined by DFT calculations, is a powerful method to determine the amide vibrations and thus the corresponding secondary structure. The calculations of the inter-atomic interactions such as TDC, leading to the splitting of the amide I band, are accurately evaluated, which is the advantage of using parameter-free, quantum chemistry methods. We extend the use of these methods to collagen (below), which also gives an idea of the size of systems that can be treated with DFT and, therefore, the range of secondary structures that could be studied with DFT, instead of using force-field methods.

III/3/ Collagen: a model for hydrated triple helical protein

III/3/1 Structural information

Collagen is the most abundant protein in mammals. It is tough and inextensible, with great tensile strength, and is the main component of cartilage, ligaments and tendons, bone and teeth. In nature, thirteen types of collagen are found and type I is the most abundant collagen of the human body. Type I collagen is constituted almost entirely by three chains of amino acids, each chain having a repeating (Gly-X-Y) sequence where Gly is the amino acid glycine, and X and Y are respectively proline and hydroxyproline. These three chains are wound together in a tight triple helix through hydrogen bonds perpendicular to the helix axis.

The conformation of the backbone of each strand of the collagen molecule is close to that of the left-handed helices PG-II and PP-II (studied above), super-coiled into right-handed triple-helices. The first synthesis of dried synthetic polypeptides (Pro-Pro-Gly)$_{10}$ was reported by Sakakibara et al.[51]. It contains approximately one water molecule per PPG triplet, equivalent to 7g water/100g completely dried (PPG)$_{10}$. Figure 18 and Figure 19 illustrate the triple helical structure of the dried (PPG)$_{10}$. The most tightly bound water molecules form hydrogen-bonded bridges between carbonyl groups of two different helices. The triple helical arrangement is favoured by the

28

presence of glycine at every third residue, reducing the steric hindrance and providing an inter-chain hydrogen bond perpendicular to the helix axis.
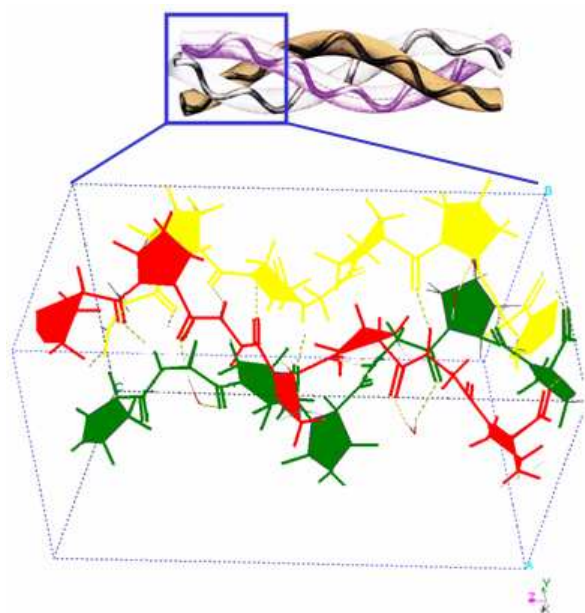
Figure 18

Structure of low hydrated $(PPG)_{10}$ [52]. The triple helices are represented with different colours.
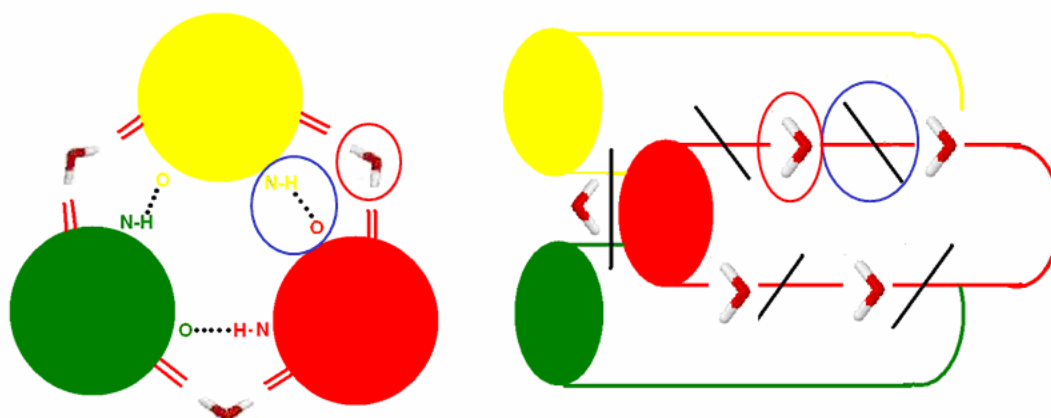
Figure 19

Triple helices seen from the top. First water shell is represented and the inter helices hydrogen bonds are perpendicular to the helices axis. The red circles represent water molecules at the first hydration level attached to the protein by hydrogen bonds. Blue circles represent inter-helices hydrogen bonds.

III/3/2/ Vibrational spectroscopy

INS has been used to measure the vibrational modes of (Pro-Pro-Gly)$_{10}$ polypeptide [39] (Figure 20).
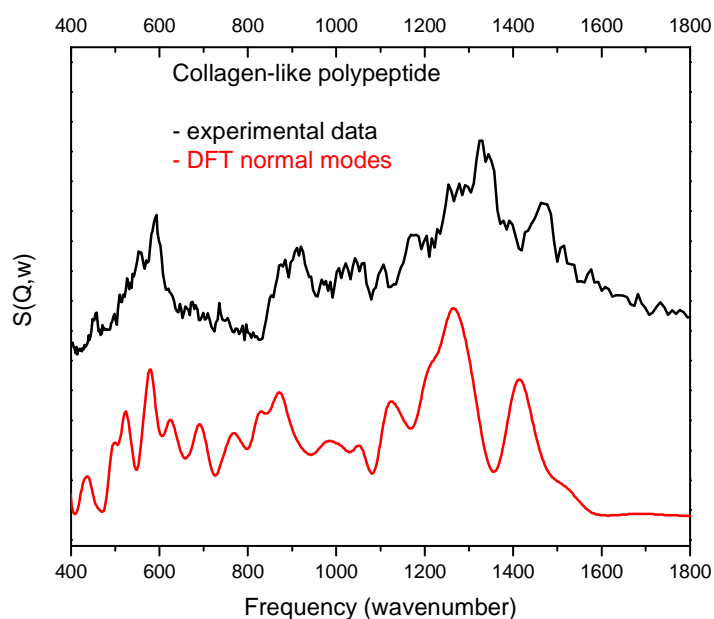
Figure 20

Comparison of measured [39] (top) and calculated (bottom) INS spectra of (Pro-Pro-Gly)$_{10}$ by normal mode methods (red curve) at the BZ centre (K=0).

As seen in the previous section, the agreement between experiment [39] and simulation is good for crystalline systems of model polypeptides of collagen (PG-II and PP-II) and therefore allows a clear evaluation of the numerical model and further analysis of helix dynamics. However, calculations on low hydrated PPG$_{10}$ triple helices reveal the importance of crystallinity in the choice of DFT-based, normal mode analysis or molecular dynamics methods. Whereas amide bands vibrations are well reproduced in the range of 400 to 1800 cm^{-1} by normal modes this is not the case for the low frequency region from 0 to 400 cm^{-1}.

In DFT methods, the ground state electronic energy is not known analytically but is reached numerically by defining the convergence criteria of the variational principle. Depending on the degree of crystallinity of the system, the DFT-determined PES of amorphous systems is poorly defined, giving non-negligible, residual forces on ions and allowing the system to explore several local minima which are close in terms of energy and configuration. Since the system is not in the global or a well-defined local

minimum of the PES, the diagonalization of the dynamical matrix gives unphysical, negative eigenfrequencies (and eigevectors). Thus, DFT-based normal mode analysis is not the method of choice to simulate dynamics of partially crystalline or amorphous systems in the low frequency range and MD, which explores local minima of the PES, is the best way to calculate the vibrational properties. MD allows a gain in precision of calculated vibrational frequencies and intensities but information on atomic polarization is lost. Understanding dynamics of amorphous systems by normal mode analysis therefore requires a low temperature structure optimized by analytical methods because empirical force fields allow a more exhaustive search for more stable minima and, otherwise, local minima can be investigated with much higher numerical precision.

III/3/3/ Amide bands: triple helical signature

As previously seen for polyglycine, the amide-I vibration is sensitive to the three dimensional form of the secondary structure. Raman and INS experiments on the collagen protein [39] has suggested that the amide-V mode is the vibrational signature of the three dimensional form of the tertiary structure.

This vibration and thereby the intensity of this band are principally due to the motion of the peptide hydrogen of glycine. This amide band appears at 707 cm^{-1} for PG-II [39] and normal mode analysis of PG-II gives six normal modes at 702, 709, 755, 758 cm^{-1}. From INS experiments on PPG$_{10}$ it has been proposed that the effect of triple-helical supercoiling is to shift the amide-V mode of the Gly-Pro linkage, together with skeletal deformation, downward to the region of 590 cm^{-1}. The authors suggested that this mode is limited to one band in the range from 550 to 590 cm^{-1} and accurate DFT-based calculations of the INS spectra allow more precise information about this type of vibration to be obtained.

While the calculated spectrum for PPG shows a strong peak at ~600 cm^{-1}, in good agreement with experiment, analysing the vibrations involving the peptide hydrogen show that these range from 580 to 780 cm^{-1} (Figure 21, bottom panel). In addition, many other hydrogens, that contribute strongly to the INS intensity, are involved in the same frequency range. The amplitude of motion of the peptide hydrogen as a

function of frequency in the frequency range 580 to 780 cm^{-1} suggests that four peaks centred at 590, 620, 693 and 760 cm^{-1} in the calculated INS spectra (Figure 20) could all be assigned to the amide-V bands. These peaks were not resolved experimentally due to the poor, sample-limited resolution of this partially amorphous system, but the decaying intensity from 650 to 800 cm^{-1} in the measured signal is consistent with the calculated spectrum.
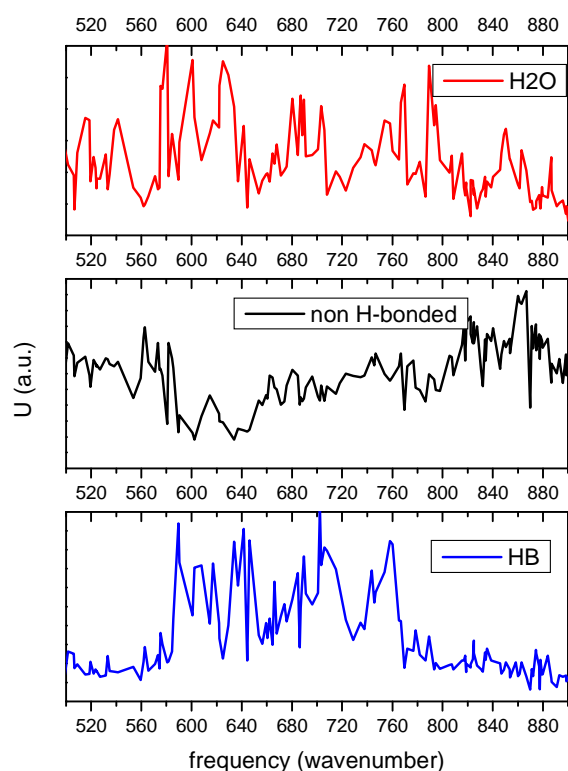
Figure 21

Amplitude of the hydrogen motion belonging to water (upper curve), to amino-acid (middle) and to peptide plane (bottom curve) as a function of the vibrational frequency.

III/3/4/ Low frequency phonons, breathing modes

The relationship between the molecular structure of collagen and its biological function, which depends on low frequency vibrations, was the subject of a number previous studies [53,54]. Based on published theoretical work, Middendorf has identified three different classes of low frequency modes below 300 cm^{-1}, the acoustic phonons (below 100 cm^{-1}), the torsion of the peptide backbone (100 to 200

cm^{-1}) and skeletal deformation (200 to 300 cm^{-1}). A tentative of assignment of low frequency vibrational modes has been performed by Bernev et al. [55]. They focused on the very low-frequency region below 100 cm^{-1} and showed that particular modes arising from INS spectra at 17, 27 and 40 cm^{-1} may represent internal breathing modes of collagen.

Due to the low cristallinity of the low hydrated structure of (Pro-Pro-Gly)$_{10}$, INS, normal mode calculated spectra are plagued by negative frequencies and don't allow evaluation of specific modes. Molecular dynamics calculations at 50K reproduce better the INS experimental profile, showing three distinct vibrational bands in the vDOS (Figure 22) which could be the three classes of vibrations proposed by Middendorff. The triple helices axis being oriented along Z, a spatial projection of the partial vDOS relative to the skeleton of the amino-acid chains can help for a better understanding of these low frequency INS bands. Figure 23 shows two regions corresponding to the ranges 0-120 cm^{-1} and 120-360 cm^{-1}, the latter representing mostly isotropic motions of the triple helices, while the former presents a radially polarized vDOS with pronounced peaks at 40 and 90 cm^{-1}.
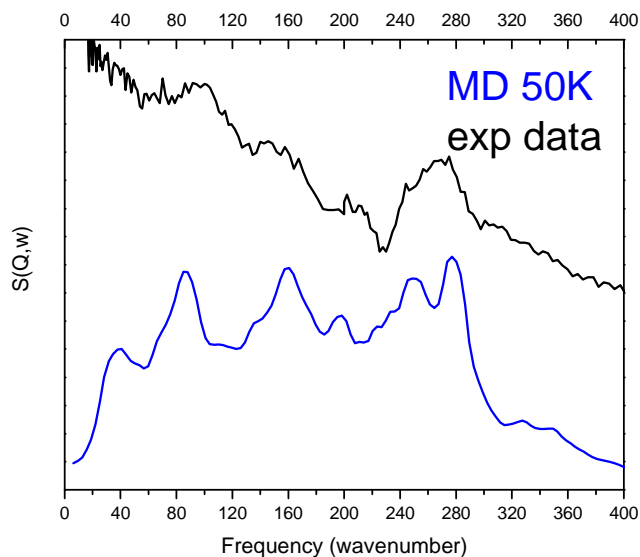
Figure 22

Comparison of measured INS spectra [39] (top) and calculated vDOS of the hydrogens of the (Pro-Pro-Gly)$_{10}$ low hydrated structure by molecular dynamics (NVT) simulations.
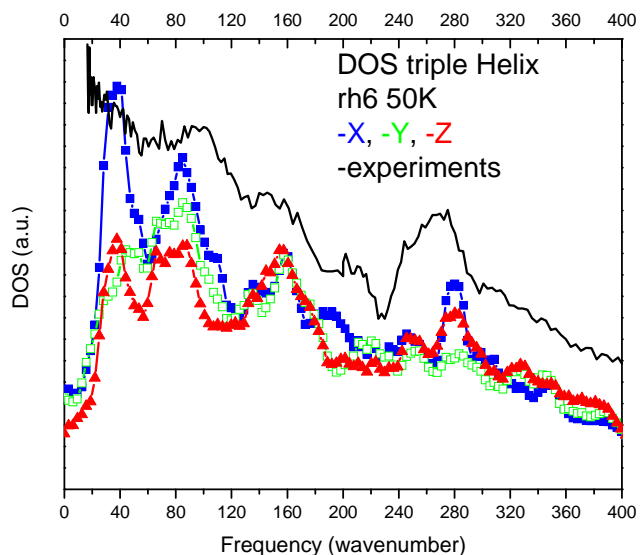
Figure 23

Measured (top) and spatial projections (on X: blue, filled squares; Y: green, open squares; Z: red, filled triangles) of the density of states of atoms involved in the triple helical chains from MD simulation at 50K. The initial structure is the lowest hydrated (PPG)$_{10}$ with a relative humidity of 7%. The measured spectrum is represented by the black curve.

Nevertheless, from MD simulations the loss of specific information relating to normal modes requires normal mode analysis to be performed using an empirical force field. The model structure was optimized using the CHARMM force field and contains 34 water molecules. In Figure 24, the low frequency INS spectrum is reasonably well-reproduced, allowing an atomistic picture of the typical vibrational modes of the triple helices to be explored.
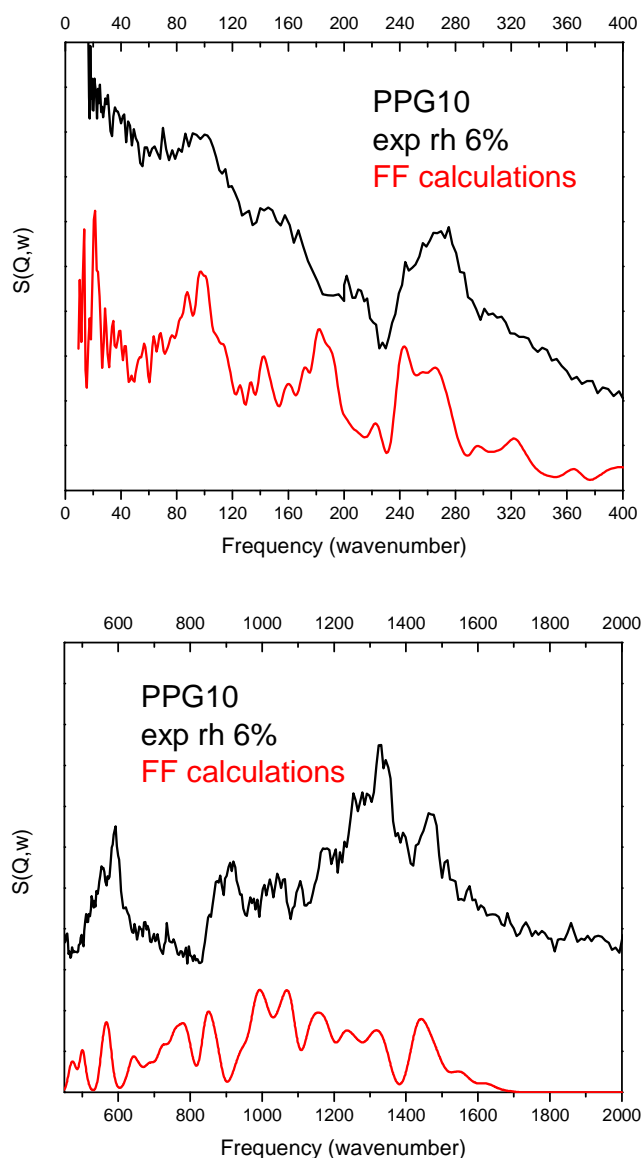
Figure 24

Comparison of measured (upper curves in black) and calculated INS spectra of (Pro-Pro-Gly)$_{10}$ by normal mode methods (lower curves in red). The low frequency region is represented in the upper panel

A more complete assignment of the vibrational modes is complicated by the large number of atoms and therefore modes. Projection operators (in addition to simple projections on z and r) can be defined to extract the nature of vibrational bands (Figure 26). The equations of the projection operators are defined below and we deal with Nt triples of residues/beads in the plane instead of with pairs where $\vec{d}_i^{\,t}$ is the

displacement vector and $\vec{r}_i^{\,t}$ the position of the i-th atom in the i-th triple and $\vec{r}_{CM}^{\,t}$ is the center of mass of the t-th triple: $\displaystyle \vec{r}_{CM}^{\,t} = \sum_i^3 \frac{m_i^t . \vec{r}_i^{\,t}}{m_i^t}$

Pure rotation (around z-axis), $\displaystyle PR = \frac{1}{N_t} \left| \sum_t^{N_t} \frac{1}{3.\max\left(\left|\vec{d}_1^{\,t}\right|,\left|\vec{d}_2^{\,t}\right|,\left|\vec{d}_3^{\,t}\right|\right)} \sum_i^3 \left(\frac{\vec{r}_i^{\,t} - \vec{r}_{CM}^{\,t}}{\left|\vec{r}_i^{\,t} - \vec{r}_{CM}^{\,t}\right|} \times \vec{e}_z \right).\vec{d}_i^{\,t} \right|$

Twisting rotation, $\displaystyle TR = \frac{1}{N_t} \sum_t^{N_t} \frac{1}{3.\max\left(\left|\vec{d}_1^{\,t}\right|,\left|\vec{d}_2^{\,t}\right|,\left|\vec{d}_3^{\,t}\right|\right)} \sum_i^3 \left| \left(\frac{\vec{r}_i^{\,t} - \vec{r}_{CM}^{\,t}}{\left|\vec{r}_i^{\,t} - \vec{r}_{CM}^{\,t}\right|} \times \vec{e}_z \right).\vec{d}_i^{\,t} \right| - PR$

Breathing $\displaystyle B = \frac{1}{N_t} \left| \sum_t^{N_t} \frac{1}{3.\max\left(\left|\vec{d}_1^{\,t}\right|,\left|\vec{d}_2^{\,t}\right|,\left|\vec{d}_3^{\,t}\right|\right)} \sum_i^3 \left(\frac{\vec{r}_i^{\,t} - \vec{r}_{CM}^{\,t}}{\left|\vec{r}_i^{\,t} - \vec{r}_{CM}^{\,t}\right|} \right).\vec{d}_i^{\,t} \right|$

Participation ratio $\displaystyle PR = \frac{1}{N_t} \frac{\left[\sum_{i=1}^{3N} \left| \vec{e}_i(\omega) \right|^2 \right]^2}{\sum_{i=1}^{3N} \left| \vec{e}_i(\omega) \right|^4}$ (ω, eigenvalue and $\mathbf{e}_i$, eigenvector)
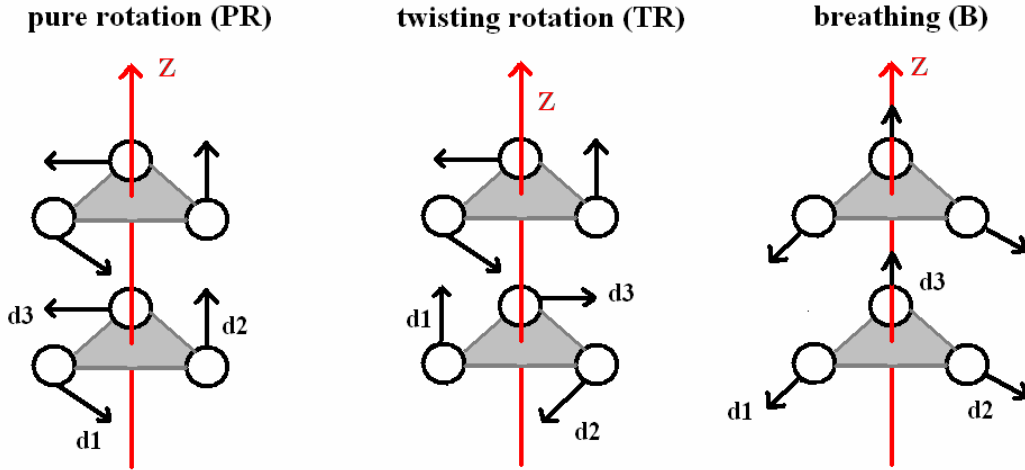
Figure 25

Bead motion corresponding to the projection equations defined above
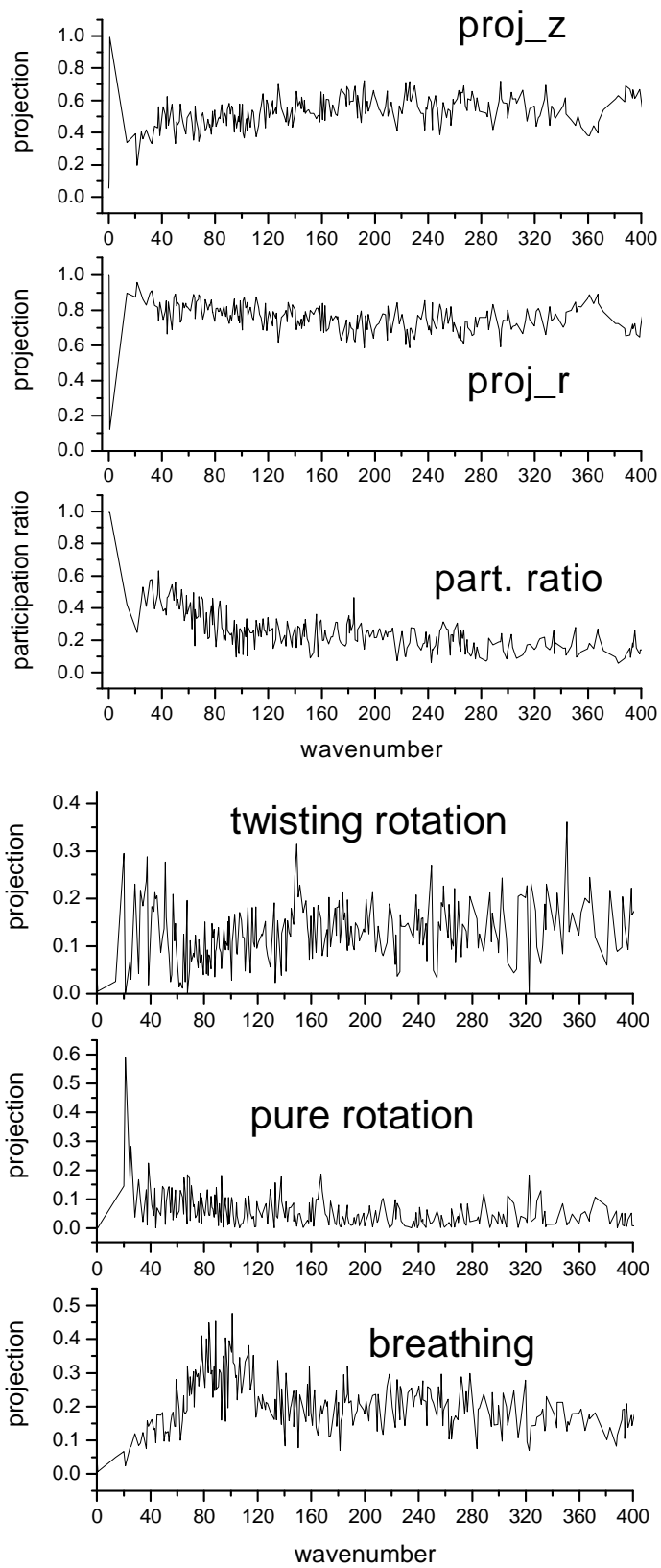
Figure 26

Projection anlaysis of eigenvectors

The radial and axial projections are broadly consistent with the DFT MD analysis (Figure 22), showing stronger radial character at low frequency and a growing axial character with increasing frequency. Only acoustic phonons are significantly delocalised and, above 100 cm^{-1}, the participation ratio reaches an approximately constant value. The torsional character of modes is pronounced between 0 and 60 cm^{-1} and at 140 cm^{-1}. Modes with breathing character (out-of-phase radial motion of the three amino acids which belong to the same plane perpendicular to the triple helices axis) are most evident between 60 and 120 cm^{-1}, with a maximum at ~100 cm^{-1}. This should correspond to the first broad INS band seen in Figure 24.

The extent of delocalization and polarization is illustrated by displacement vector plots for particular modes. Figure 27 shows three modes with pronounced radial character due to twisting rotation (left panel), due to pure rotation of the helices (middle panel) and due to breathing character (right panel). The third mode has the strongest "breathing" polarization of all modes.
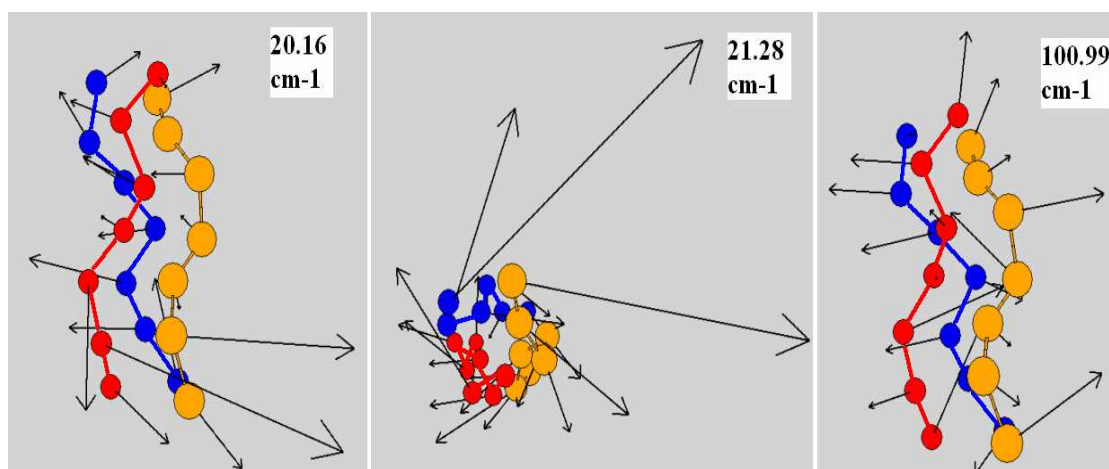
Figure 27
Some typical bead eigenvectors

III/3/5/ Temperature and anharmonicity

Benefiting from not imposing the harmonic approximation and allowing finite temperature to be investigated, INS spectra of collagen calculated with MD methods provide a clear insight into the temperature dependence of the anharmonicity of

protein dynamics (Figure 28). An anharmonic shift of 20 cm^{-1} is present at low frequency (below 300 cm^{-1}) for each of the three broad vibrational bands.
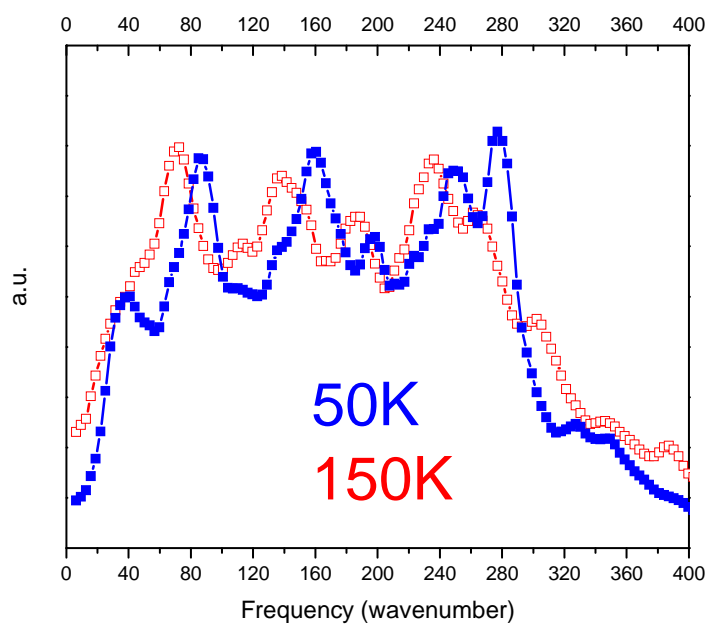
Figure 28

Comparison of calculated vDOS of hydrogens of (Pro-Pro-Gly)$_{10}$ at 50K (blue, filled squares) and 150K (red, open squares).

III/3/6/ Water shell dynamics

Finally, via the vibrational DOS of the water hydrogen atoms, we focus on the water dynamics surrounding the collagen. Depending on the first hydrogen-bonded neighbours, two different water shells can be differentiated: tightly bound to the protein, characteristic of the first hydration level, and an inter-helical water shell, characteristic of an upper hydration level (Figure 29). The first hydration shell can also include water molecules bridging between different helices of the triple helix.
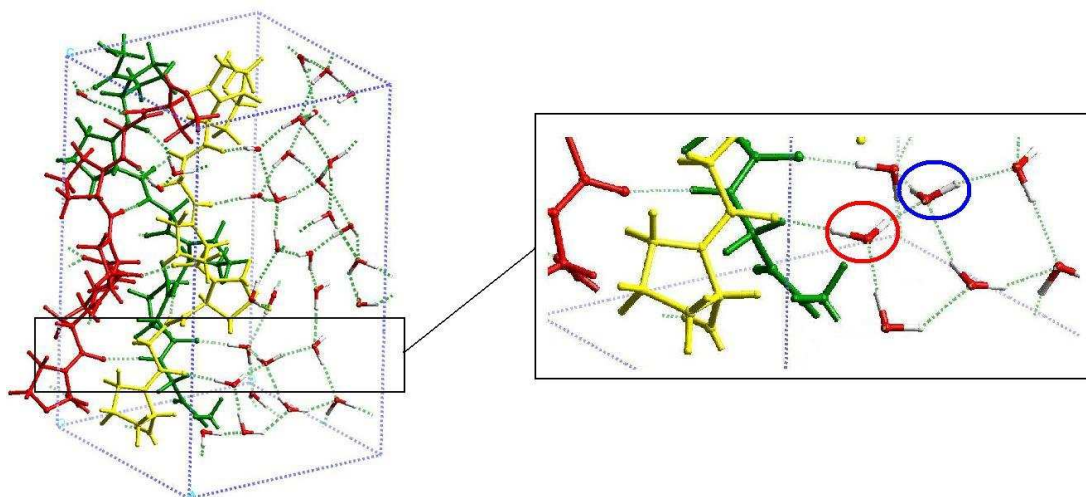
Figure 29

Left picture: structure of high hydrated (PPG)₁₀. A zoom of the interface protein-water is represented on the right figure. The red circle represents water in the first hydration shell sharing hydrogen bond with free oxygen of the protein. The blue circle represents water in the second hydration shell sharing four hydrogen bonds with other water molecules, ressembling bulk water.

The vDOS (Figure 30) indicates that both tightly bound and interhelical water have similar librational dynamics. Both types of water have modes in the frequency ranges 0-300 cm^{-1}, 600-1000 cm^{-1} and around 1600 cm^{-1}. The lowest frequency band matches that of the triple helix, while the intermediate frequency excitations suggest coupling between water and the amide-V vibrations (see Figure 21). The main difference between the two types of water is at low frequency, the inter-helical water having a stronger vDOS since it can support collective excitations of extended hydrogen-bonded water clusters.
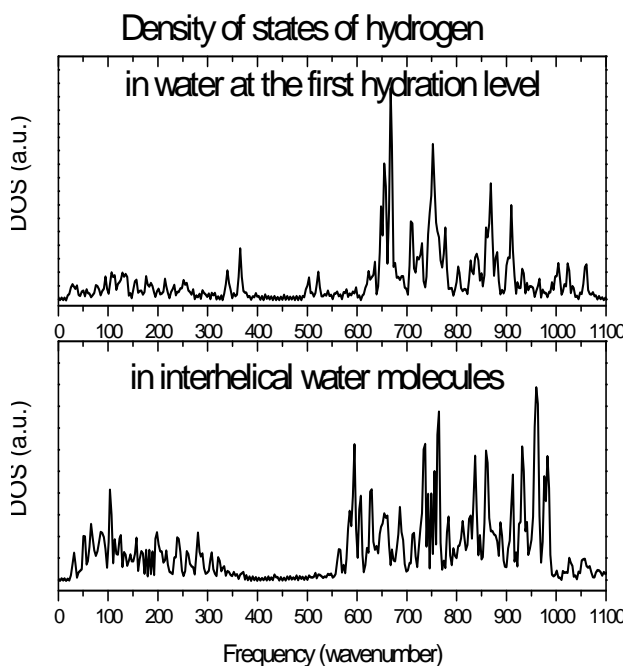
Figure 30

Vibrational density of states (vDOS) of water's hydrogen located in the first hydration level (upper panel) and in interhelical water shell (bottom panel).

III/4/ Conclusion

In this chapter we have investigated 4 structures of PPTA. INS has clearly been able to confirm what is currently regarded, from diffraction measurements, as the correct structure. Even though the resolution of the experimental data is limited, in this case by the sample, the fact that the whole spectral profile is directly determined from the calculated normal modes adds confidence to this result. The vibrations that are most sensitive to the packing of 2-D sheets are in fact amide-V vibrations but we emphasise that the whole spectral profile is used to identify the structure. In the case of PPTA, it is also worth highlighting that the structural feature that is being probed is the inter-sheet packing, whereas, generally, in proteins the goal is to identify and quantify the presence of secondary structures.

We have also studied PG-I, PG-II and PP-II. The calculated INS profiles ressemble closely the calculated profiles. The analysis of the amide-I band, experimentally, shows trends that are consistent with the established markers for beta sheets and helices. Numerically, the results are similar, but the precision is of the order of 10-20 cm^{-1}, which would seem to be insufficient to exploit IR data on complex proteins.

However, the splitting of this band is well reproduced and confirms the accuracy of DFT methods to calculate inter-atomic interactions such as TDC coupling. In this study, the calculations serve to characterize the purity of the amide-I modes and one sees clearly that there is significant variation between the displacement vectors of the amide groups in the helical structures of PG and PP. One important point to emphasize is that the experimental data and calculations performed here relate to the solid state, which imposes constraints on the secondary structures. The structure-dynamics relation may be more precise in the more flexible, less constrained environment of a liquid in which "ideal" secondary structures can exist. In our simulations, the inter-helix interactions could be reduced by using a bigger unit cell, which remains to be investigated.

Collagen has been investigated in detail with normal mode and MD methods. It is the biggest system (~350 atoms) that we have tackled with DFT methods. Reasonable agreement is obtained between the DFT-based normal mode calculation and the INS data, although the amorphous water results in a numerically, poorly-defined minimum of the PES and therefore a significant number of negative frequencies.

Via the amide-V vibration, the structure-dynamics relation has been established to investigate the signature of tertiary structure formation. Normal mode calculations suggest several bands which have not been seen experimentally due to the poor resolution of the INS spectra.

The low frequency spectrum from the MD simulation is therefore analysed, via the Fourier transform of the time dependence of geometrical quantities that have been defined to reveal different types of collective modes. This analysis is not very direct and a highly-converged, force field-based, normal mode calculation has also been performed. At high frequency (> 500 cm^{-1}) the agreement between calculation and experiment is not so good, indicating a lack of precision in certain force constants (see Figure 24).

But at low frequency, where the collective modes depend on the average of many force constants, the agreement is much better. Although, in principle, exploiting displacement vectors is easier than MD trajectories, when there are ~1000 degrees of

freedom, this is not so clear. A number of projection operators have been defined, which help to define the character of modes as a function of frequency. In particular, breathing modes in which the three helices move away from each other, stretching the inter-helix hydrogen bonds, are found to have a characteristic frequency of 90 +/- 10 cm^{-1}.

Via the kinetic energy of the atoms, DFT-MD simulations give access to the temperature dependence of the protein dynamics and reveal significant anharmonicity (around 10 cm^{-1}) of three low frequency bands between 0 and 400 cm^{-1}. This method also allows us to investigate different types of water dynamics depending on the water shell surrounding the protein.

Technically the key result from this work on biopolymers is that DFT-based, normal mode calculations work well on crystalline systems, but they are prone to negative frequencies in partially crystalline/amorphous systems, since a numerically, well-defined minimum of the PES can not be found. Force field methods, which depend on the numerical evaluation of analytical expressions do not have this problem, but they are based on simpler forms of interactions that are parameterised. The limitations of this parameterisation are seen at high frequency, but at low frequency, dynamics with a more collective character *appear* to be more accurately described. We pursue these ideas in the final chapter on DNA for which the large number of atoms and calculations to be performed exclude the use of DFT.

Chapter 4

DNA, the holy-grail of bio-simulation

IV/1/ Introduction

Usually called the molecule of heredity, DNA contains the genetic instructions specifying the biological development of all cellular forms of life, and many viruses. Understanding the function of biological molecules has evolved from being structure-based to including a knowledge of molecular dynamics. In the case of DNA, a key point of interest is base-pair opening as such dynamics are expected to play a key role in replication, transcription and denaturation [56,57,58]. Otherwise DNA presents five, well-defined structures and dynamics must play a key role in structural transitions. Since biological activity takes place at or close to room temperature, the focus tends to be on low frequency vibrations of biological molecules.

The dynamics of DNA has attracted considerable attention. Raman spectroscopy has been used to study low frequency dynamics of A- and B-DNA [59,60] in solution and in liquid crystalline forms as a function of temperature. Peaks and broad features can only be described as corresponding to localized or delocalized modes, the latter being intra- or inter-helical. Optical spectroscopy is limited to k~0, but dispersion associated with delocalized modes is typically measured by coherent, inelastic neutron scattering (INS). For large systems, diffusion at small wave vectors should be measured but kinematic limitations make this impossible for INS and a strongly dispersive mode was measured in the vicinity (~2 A-1) of the base pair Bragg peak [61]. The kinematic problem does not apply to inelastic X-ray scattering since the X-ray energy is much greater than the energy transferred to the sample and the same dispersive mode has been measured close to the gamma point in a range of DNA samples [62,63]. The combined result of these INS and IXS measurements is an acoustic-type phonon that has a maximum frequency of ~12 meV at ~1 A-1, a minimum of 2 meV at ~2 A-1 and at higher Q the phonon is strongly damped and can no longer be measured.

Despite the obvious importance of DNA, no clear description or understanding of these modes exists. Simplified models using beads to represent nucleotides have been developed [64,65,66,67,68]. Imposing C2 symmetry between strands and helical S10

symmetry (for B-DNA) results in 3 in-phase and 3 out-of-phase modes, where the degrees of freedom are longitudinal (z), radial (r) and torsional (theta or phi). Base-pair opening is described by the out-of-phase, radial mode. Since inter-atomic force constants cannot be used in a bead model, interactions have to be refined for the beads and this resulted in a fully delocalized, base pair-opening mode at ~100 cm^{-1}. The relevance of a such a model to DNA composed of random base-pair sequences and surrounded by partially crystalline water is discussed below.

IV/2/ Numerical details and vibrational spectroscopy

Atomistic force fields and molecular dynamics simulations for biological molecules are widely used to understand the behaviour of such complex systems. For DNA, the CHARMM force field and MD code has been shown to reproduce average, structural properties of different forms of DNA, but it has not been applied to study phonons and molecular vibrations. We have done precisely this to obtain an atomistic picture of the vibrational modes of DNA. In order to compare our B-DNA model with the INS and IXS phonon data, we have set up a phonon calculation for arbitrarily large systems, using the method and approach embodied in the PHONON code [69]. Our code calculates spectral observables for IXS and INS, for coherent and incoherent scattering in the second case. This allows us to compare calculated vibrations over the whole spectral range with new, incoherent INS data.

An initial B-DNA structure was obtained from the Protein Data Bank. This structure was adapted slightly to give a rectangular, periodic model with dimensions $a=b=32.3Å$ and $c=33.5Å$ where c is parallel with the helix axis. The 10 base-pair sequence is CTCTGCTACT (for 1 strand) and the simulation box was filled with almost 1000 water molecules giving a density of 1.14 g/cm^3. MD (NVE) simulations at 300K, followed by a quench were used to generate the equilibrium structure. For such large systems, the Hessian has to be determined by finite displacements and this matrix was given as input to our phonon code. Diagonalising the dynamical matrix for wave-vectors along the z (helical) axis gives the dispersion curves shown in Figure 31, left panel. We note that the acoustic phonons have a maximum frequency of ~1 meV at the Brillouin Zone boundary. At higher frequencies, we see optic modes with limited dispersion. However calculating the spectral intensities for coherent IXS and

INS for wave vectors up to 3A-1 gives a dispersion curve (Figure 31, right panel), which agrees well with the experimental data in terms of frequency and width. The experimental signal is not therefore a direct measurement of the acoustic phonon but it is the projection of spectral intensity over a large number of mainly optic modes. The sound velocity derived from Figure 31 at low q and over an extended q range (< 0.5A-1) is ~3000 m/s. At the gamma point the optic modes start at 14 cm^{-1} and there are ~1000 modes in the first 100 cm^{-1}. This result is consistent with the observation by Raman spectroscopy of an optic mode at 16 cm^{-1} for B-DNA. Since we do not calculate Raman intensity, we cannot exclude the possibility of the spectral profile having a maximum at higher frequency, but the dispersion curves show that the optic mode cannot occur at lower frequency.
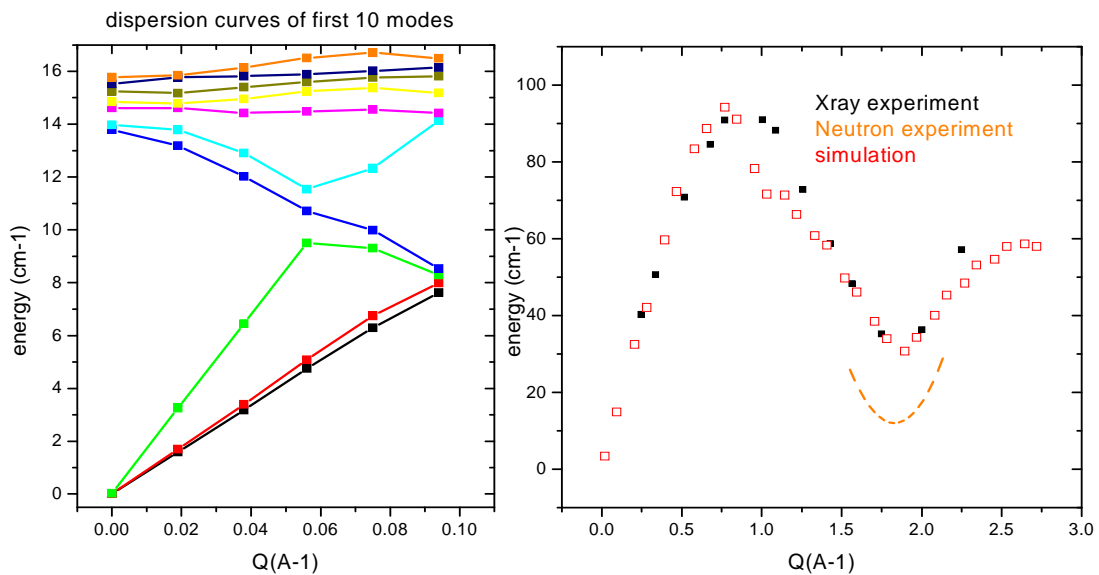
Figure 31

Comparison of measured (filled squares and dotted line) and calculated (open squares) dispersion curves (right panel) and a zoom presenting the energy limit of the acoustic modes (left panel)

Figure 32 and Figure 33 show measured and calculated incoherent INS spectra up to 2000 cm^{-1}.
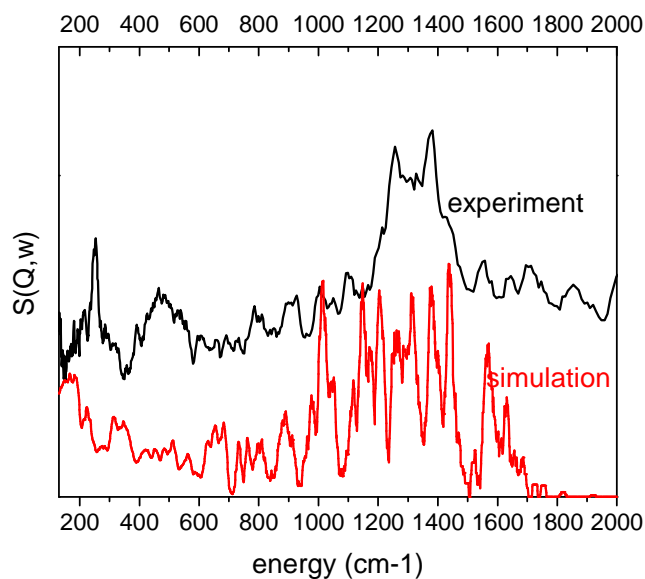
Figure 32

Measured on TOSCA instrument (upper curve) and calculated incoherent INS spectra of hydrated DNA
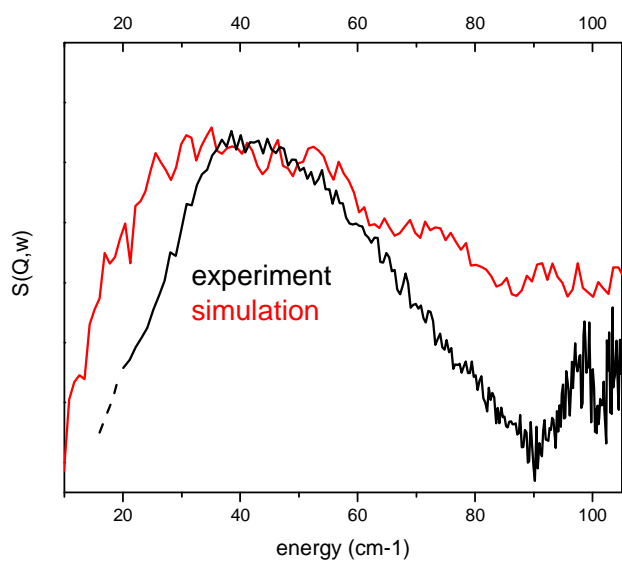
Figure 33

Measured on IN4 instrument (lower curve) and calculated incoherent INS spectra of hydrated DNA

The agreement between experiment and simulation is reasonable, allowing the principle features of the higher frequency spectrum to be assigned, for example the methyl group libration at ~250 cm^{-1} and the so-called sugar band at 1200-1500 cm^{-1} [70]. The limitations of the CHARMM parametrisation for nucleotides is seen in the 50 cm^{-1} discrepancy for the methyl group libration (Figure 34) and the strong band at

1000 cm^{-1} in the calculated spectrum, which inspection of the eigenvectors shows to be related mainly to the sugar and the methyl group of the thymine (Figure 35).
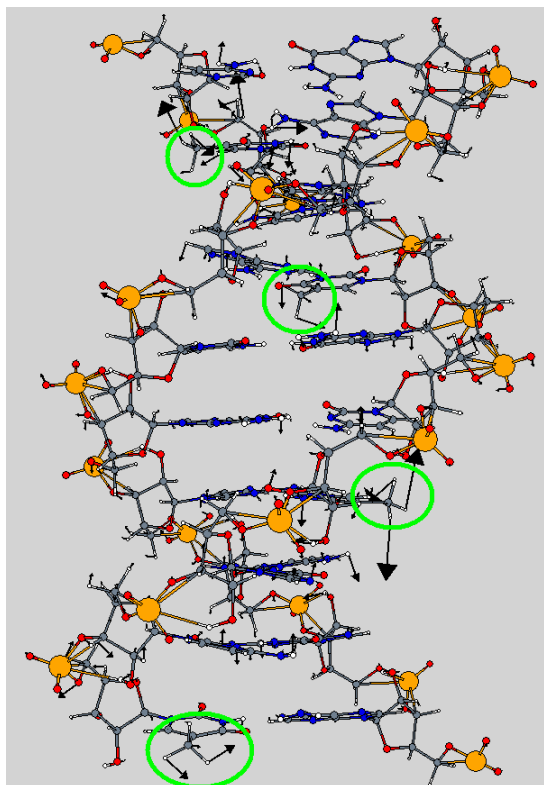
Figure 34

Eigenvectors corresponding to methyl rotation of thymine (in light green) at ω = 228 cm^{-1}
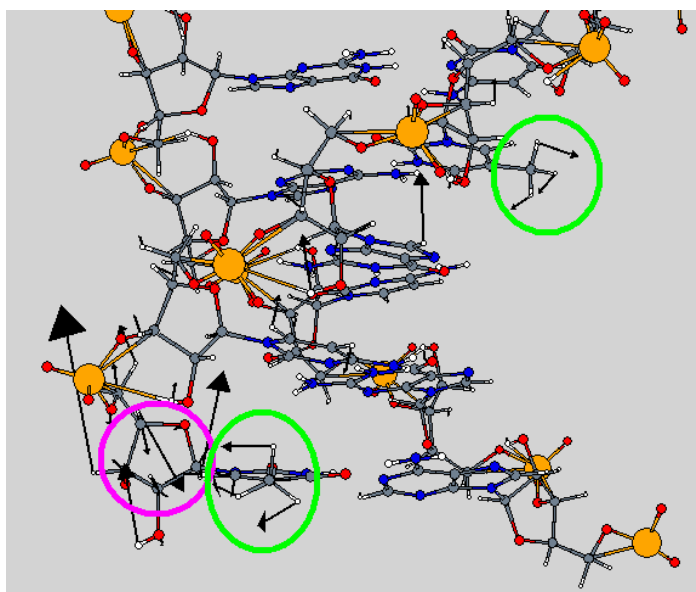
Figure 35

Eigenvectors corresponding to the coupling between sugar (pink) and CH3 deformation of thymine (light green) at ω = 1009 cm^{-1}

A more complete assignment of the vibrational modes is complicated by the large number of atoms and therefore modes. In addition we have taken a random sequence of base pairs which, with the water, tends to smear out any particular vibrational character. Projection operators (in addition to simple projections on to z and r) can be defined to extract the nature of vibrational bands as follows, according to the degrees of freedom of a bead model in which the beads are nucleotides.

$\vec{d_1}$ and $\vec{d_2}$ are the net displacement vectors of beads, defined as nucleotides, and $\vec{r}$ is the inter-bead vector.

$$\text{Pure rotation } PR = \frac{1}{N_p} \left| \sum_p^{Np} \frac{\left(\frac{\vec{r}}{|\vec{r}|} \times \frac{\vec{z}}{|\vec{z}|}\right).(\vec{d_1} - \vec{d_2})}{2.\max(|\vec{d_1}|,|\vec{d_2}|)} \right|$$

$$\text{Twisting rotation } TR = \frac{1}{N_p} \sum_p^{Np} \frac{\left|\left(\frac{\vec{r}}{|\vec{r}|} \times \frac{\vec{z}}{|\vec{z}|}\right).(\vec{d_1} - \vec{d_2})\right|}{2.\max(|\vec{d_1}|,|\vec{d_2}|)} - PR$$

$$\text{Breathing } B = \frac{1}{N_p} \sum_p^{Np} \frac{\left|\frac{\vec{r}}{|\vec{r}|}.(\vec{d_1} - \vec{d_2})\right|}{2.\max(|\vec{d_1}|,|\vec{d_2}|)}$$

$$\text{Out of axis shifting } S = \frac{1}{N_p} \sum_p^{Np} \frac{\left|\frac{\vec{r}}{|\vec{r}|}.(\vec{d_1} + \vec{d_2})\right|}{2.\max(|\vec{d_1}|,|\vec{d_2}|)}$$

$$\text{Out of plane swinging } OS = \frac{1}{N_p} \sum_p^{Np} \frac{\left|\frac{\vec{z}}{|\vec{z}|}.(\vec{d_1} - \vec{d_2})\right|}{2.\max(|\vec{d_1}|,|\vec{d_2}|)}$$

$$\text{Plane swinging } PS = \frac{1}{N_p} \sum_p^{Np} \frac{\left|\frac{\vec{z}}{|\vec{z}|}.(\vec{d_1} + \vec{d_2})\right|}{2.\max(|\vec{d_1}|,|\vec{d_2}|)}$$

Participation ratio $PR = \dfrac{1}{N} \dfrac{\left[\sum\limits_{i=1}^{Np} \left| \vec{e}_i(\omega) \right|^2 \right]^2}{\sum\limits_{i=1}^{Np} \left| \vec{e}_i(\omega) \right|^4}$ (ω, eigenvalue and $\mathbf{e}_i$, eigenvector)
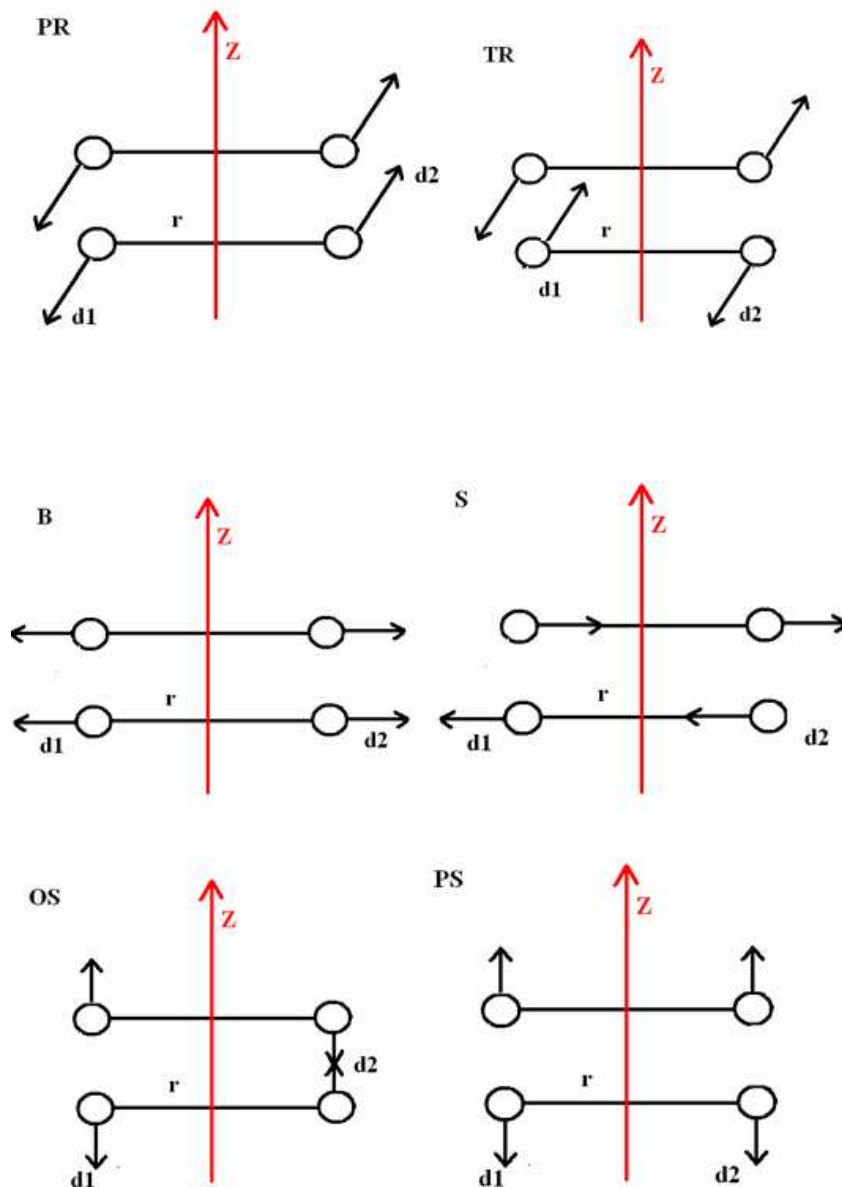
Figure 36

Several types of bead motions corresponding to the typical projection operators defined above.

Modes tend to be delocalized over the first 50 cm^{-1} and at higher frequency (upto 300 cm^{-1}) there is a significant decay in the participating ratio. Modes with breathing character (out-of-phase radial motion of the nucleotides) peak at ~120 cm^{-1}. At this frequency vibrational modes are already significantly localized, which means there is

no coherent base pair opening over the whole helix. The in-phase, radial motion of the nucleotides is strongest at low frequency (< 100 cm^{-1}) since these modes do not involve stretching of the inter-base hydrogen bonds. In accordance with these "breathing" projections, the simple radial projection shows maxima at (>) 120 cm^{-1} and 50 cm^{-1}. The axial projection is almost a direct inversion of the radial projection. The out-of-plane rocking of the nucleotides is pronounced at ~25 cm^{-1}. The torsional character of modes is only marginally pronounced at 50 cm^{-1}.
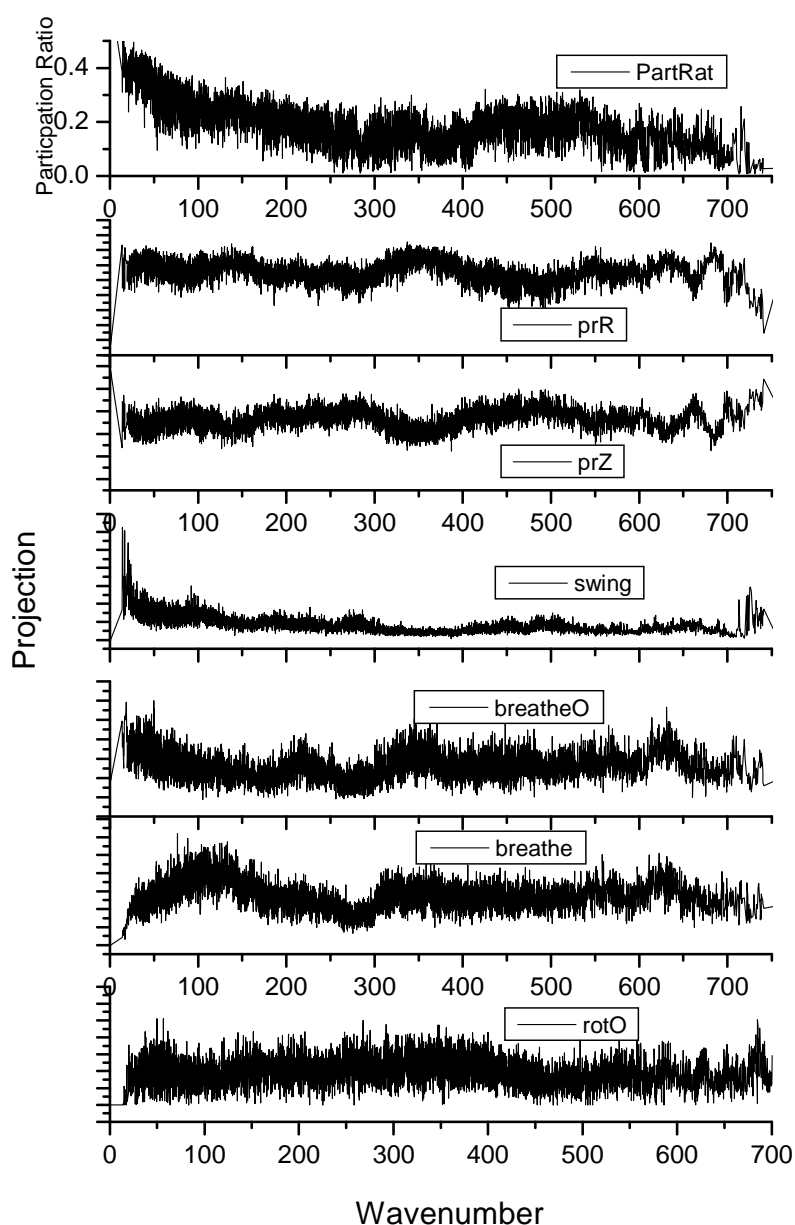
Figure 37

projection anlaysis of eigenvectors

The extent of delocalization and polarization is illustrated by displacement vector plots for particular modes. Figure 38 shows three modes with pronounced axial character (left panel), with radial character due to in-phase motion of nucleotide pairs (middle panel) and with radial breathing character (right panel). The third mode has the strongest "breathing" polarization of all modes. The PaR for this mode is about half of the PaR for the lower frequency modes shown in Figure 37 and, accordingly, only about four nucleotide pairs have significant displacements.
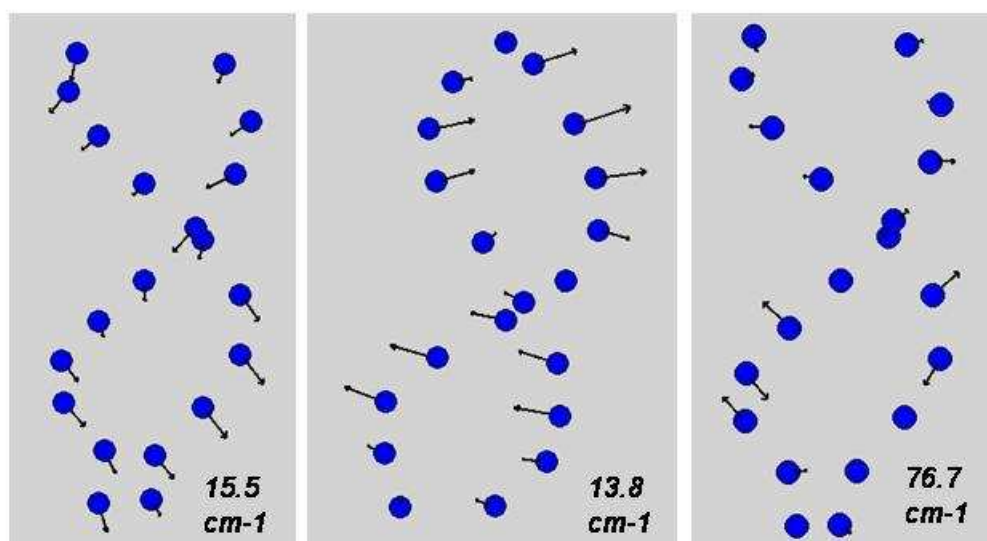
Figure 38
Some typical bead eigenvectors

IV/3/ Conclusion

In this chapter we have investigated the dynamics of a B-DNA structure in its natural environment of water and counter-ions by comparing new experimental INS data with normal mode calculations. The large number of atoms in such structure prohibits the use of DFT-based calculations and FF methods combined with large scale phonon calculations were employed to overcome this problem. These calculations allow us to describe, with reasonable reliability, the vibrational modes over the whole complex spectrum of DNA. However, the limitations of the CHARMM parametrisation for nucleotides are clearly seen at 250 and 1000 cm^{-1} which are related to methyl group and sugar vibrations [annexe F/3/]. At low frequency the incoherent INS spectrum is feature-less, but the dispersion curved measured by IXS and INS is accurately

reproduced. Analysing the low frequency vibrations, a band of modes centred at 100 cm^{-1} is found with strong radial breathing character and partial localisation over about four of ten base pairs. These breathing modes are considered to play a key role in the base-pair opening since they may be related to the dynamics of the DNA bubbles [71]. The more reliable calculation of low frequency modes using CHARMM and breathing modes corresponding to the opening of hydrogen bonds at ~100 cm^{-1} is common to DNA and collagen (chapter 3).

Chapter 5

Conclusion

The last century has seen the triumph of experimental physics techniques to solve fundamental biological mechanisms. In 1953, the discovery of the DNA double helix structure with X-rays diffraction techniques has become a cornerstone of our comprehension of information transfer in living material. However, understanding the function of biological molecules has evolved from being structure-based to include the role of dynamics. A wide range of experimental techniques now exist to investigate molecular dynamics but the interpretation of data is generally less direct than that of diffraction data. Thus, with the continuous growth of computing resources, in terms of hardware and software, computer simulations are playing an important role in establishing the structure-dynamics-function relationship.

The starting point for this thesis project was the use of DFT and INS to study mainly the higher frequency vibrations of molecular crystals. These vibrations are generally too high in frequency to play any role in biological activity at ambient temperature. However, from our understanding of the PES, we have seen that structure and high frequency vibrations are intimately connected and that this is the basis of the widespread use of IR spectra to quantify the presence of secondary structures in proteins. The lack of long range structural order which is characteristic of polymeric systems hinders the use of diffraction techniques to quantify the presence of secondary structures (helices and sheets) in proteins. Experimentalists overcome this problem by exploiting the intimate connection between the structure and high frequency vibrations via the IR spectra profile of the amide bands.

We have used a similar approach (DFT and INS) to determine the spectral profile of the amide-V band and therefore the packing of 2D molecular sheets in Kevlar, which is a more subtle problem than the coiling of, for example, polypeptides into helices. We have also used these methods to investigate the characteristic amide bands of simple polypeptides like polyglycine and polyproline and have seen that the trends, for example, in going from a beta sheet to a helix, are well-reproduced by DFT methods even if the absolute precision is of the order of 10-20 cm-1. The periodic DFT methods that we have used can therefore be used reliably to make a quantitative,

parameter-free link between structure and dynamics of ordered, polypeptide structures. These methods could therefore be used to make a theoretical investigation of the empirical data that is used to interpret IR spectra when many secondary structures are present and when these are composed of various peptide sequences.

Low frequency dynamics can be thermally activated at room temperature and these can be vibrational (oscillatory) or dissipative, the former being of interest in this thesis. These dynamics are involved in the structure-dynamics-function relation. SSHBs are thought to play a key role in many bio-physical processes being intermediate in stability between inter and intra-molecular interactions and being able to store a considerable amount of energy. We have studied SSHBs in a range of model systems, the function of these systems being a temperature-driven chemical reaction, namely proton transfer. By combining DFT-based, normal mode analysis and MD simulations we have seen that vibrational entropy drives proton transfer and we have been able to identify the important vibrational modes. Again the precision of DFT is sufficient to qualitatively reproduce the proton transfer effect in 3 out of 4 systems, but in the fourth system, the low temperature structure of the hydrogen bond was determined to be normal rather than short, precluding any proton transfer as a function of temperature.

A key feature of ordered/crystalline bio-molecules is that they are often surrounded by water, which is crystalline in part and amorphous in others. We have studied the low frequency dynamics in two such systems, collagen and DNA. The structure of collagen is small enough (~350 atoms) to allow us to use DFT methods, which are seen to suffer from a poorly defined minimum in the PES for the partially amorphous systems, resulting in ~5% of modes having negative frequencies in a normal mode analysis. Using FF methods which have superior numerical precision, removes the problem with negative frequencies, but introduces a parameterization of the total energy calculation. The low frequency spectrum is reasonably well-reproduced by the FF methods, but this is not true at higher frequency (> 500 cm-1) where the modes depend on, and reveal inaccuracies in fewer force constants.

In DNA we have used FF methods since the models contain more than 1000 atoms and thousands of calculations have to be performed in a normal mode analysis. As for

collagen, low frequency excitations of DNA, like the measured pseudo-acoustic dispersion, seem to be well calculated but inaccuracies appear at higher frequency. In DNA, the structure-dynamics-function that we have focused on is base-pair opening, which is related to various bio-physical processes like replication and transcription. From the normal mode analysis we have identified modes with base-pair opening character grouped around 100 cm-1. This value is consistent with theoretical, coarse-grain models which have been parameterized. In our calculations we have used the empirical FF CHARMM, but the parameters in this force field have been developed over many years to be consistent with a wide range of experimental and quantum chemical data and they have not been modified in our calculations.

The tools that we have used in this thesis are neutron scattering and DFT and FF simulation methods. Neutron scattering for diffraction is the best technique for locating hydrogen atoms, which has been crucial in revealing the proton transfer effect in SSHBs. In this thesis we have mainly exploited inelastic neutron scattering, which allows a direct comparison between measured and calculated spectra and therefore an overall evaluation of the precision of the numerical model. Numerically DFT generally offers high precision over the whole range of molecular vibrations in crystalline materials, but low frequency modes in partially disordered materials suffer essentially from a lack of numerical precision. This is an important result in checking what was regarded as the unstoppable advances of DFT towards bigger and more complex systems. The calculation of low frequency modes seems to be quite good using the CHARMM FF, where each mode depends on a large number of force constants. This result is consistent with modeling of quasi-elastic neutron scattering data with force fields which tends to have good accuracy. But the limitations of the FFs are seen at higher frequencies where modes depend on the precision of a few force constants and cancellation of errors does not occur.

This thesis has therefore established a number of clear messages for future work. The numerical precision of DFT needs to be increased to handle low frequency excitations in amorphous systems. DFT could be used in a coarse-grain approximation to treat systems like hydrated DNA, which would reduce the number of calculations to be performed and generally involve bigger forces being induced on larger objects than atoms. However, we have tried to simplify the analysis of our DNA calculations by

summing over force constants in the Hessian matrix, according to a bead model, and found that this modifies considerably the description of the vibrational modes, presumably due to the neglect of rotational degrees of freedom of the 3D beads. Mapping atomic force constants onto bead force constants is not so simple and needs to be done carefully. FF methods should use complete vibrational spectra to refine force constants so that good agreement at low frequency would not be regarded, in part, as fortuitous.

An important part of this thesis has been setting-up the wet-spinning machine at ILL. This has been done successfully and opens the door to new experimental work. Two directions have been defined. The first concerns bubble formation in DNA, which relates to the base-pair opening modes investigated in this work. Bubble fluctuations require energy-resolved diffraction to be performed and the samples will have to be taken close to the denaturation limit. This type of work is now possible since oriented DNA samples can be produced at ILL. Numerical investigation of bubbles will be based on the MD simulations that we have used to generate equilibrium structures of hydrated DNA. The simulated samples can be taken arbitrarily close in temperature to the denaturation limit and even beyond! A second theme of investigation could be studying the dynamical signature of humidity-driven structural transitions. The B-form of DNA turns into the A-form on drying. Some basic geometric information (cell and helix-pitch variation) is known about this transition, which would be enough to set-up numerical models. Quasi-elastic and inelastic data could be collected and compared with model predictions.

Summary

The purpose of this thesis was to extend recent works on structure and dynamics of hydrogen bonded crystals to model biomolecular systems and biological processes.

The tools that we have used are neutron scattering (NS) and density functional theory (DFT) and force field (FF) based simulation methods. The quantitative and parameter-free link (in the case of DFT methods) between structure and dynamics has been applied to strong hydrogen bonded crystals and bio-polymers such as collagen and DNA.

In several SSHB crystals, DFT normal modes and molecular dynamics calculations revealed the mechanism of proton transfer as being driven by low frequency phonons.

The natural extension of these methods was oriented to polymers. Due to the lack of long range order, obtaining structural information of amorphous bio-polymeric systems requires the determination of the amide bands, which are the vibrations of the peptide groups C(=O)-N-H. We have used the DFT and inelastic neutron scattering approach to determine the spectral profile of the amide-V band and therefore the packing of 2D molecular sheets in Kevlar and the signature of the tertiary structure in Collagen. Depending on the secondary structure, the trends of the amide-I band has been well reproduced for simple polypeptides chains like polyglycine and polyproline.

Water surrounding the protein is a huge subject of research. Water molecules are linked together to form different amorphous hydration shells. DFT methods are seen to suffer from a poorly defined minimum in the PES, resulting in negative frequencies in a normal mode analysis. Using force fields methods overcomes this problem but introduces a parameterization of the total energy calculation.

In DNA, the structure-dynamics-function that we have focused on is base-pair opening, which is related to various bio-physical processes like replication and transcription. We used force field methods and normal mode analysis to identify modes with base-pair opening character. The oriented DNA films for experiments

were made using the wet spinning method and the equipment was successfully installed and modernized during the thesis at ILL.

Samenvatting

Het doel van dit proefschrift is het verder uitwerken van recent werk aan structuur en dynamica van waterstofgebonden kristallen om biomoleculaire systemen en biologische processen te modeleren.

De methodes die we gebruikt hebben zijn neutronenverstrooiïng (NS) en simulatietechnieken gebaseerd op density functional theory (DFT) en force fields (FF). De kwantitatieve en parametervrije relatie (in het geval van DFT) tussen structuur en dynamica is toegepast op sterk waterstofgebonden kristallen en biopolymeren, zoals collagenen en DNA.

Uit de DFT normaal modi en moleculaire dynamica berekeningen blijkt het mechanisme van proton overdracht gestuurd te worden door fononen van lage frequentie.

De voordehand liggende uitbreiding van deze methodes was gericht op polymeren. Door het gebrek aan ordening op langere afstanden, kan structurele informatie van amorfe biopolymeer systemen alleen verkregen worden via de bepalingen van amide banden (de vibraties van peptide groepen C(=O)-N-H). We hebben de combinatie van inelastische neutronenverstrooiïng en DFT gebruikt om het spectrale profiel van de amide-V band te bepalen en daarmee de stapeling van tweedimensionale moleculaire lagen in Kevlar en tekens van de tertiaire structuur van collagenen. Afhankelijke van de secundaire structuur zijn de trends van de amide-I band goed gereproduceerd voor de eenvoudige polypeptide structuren, zoals polyglycine en polypyroline.

Water dat een eiwit omringt is een zeer belangrijk onderzoeksonderwerp. Watermoleculen vormen samen netwerken waardoor verschillende amorfe hydratie schillen ontstaan. Met DFT methodes blijkt het minimum in het potentiaal landschap vaak moeilijk te bepalen, wat tot negatieve frequenties leidt in de normaal modi analyse. Met het gebruik van FF methodes wordt dit probleem omzeild, maar wordt ook een parametrisatie geintroduceerd bij de berekening van de totale energie.

In DNA hebben we de structuur-dynamica relatie onderzocht van het openen van baseparen, dat gerelateerd is aan verscheidene biofysische processen, zoals replicatie en transcriptie. We hebben FF methodes en normaal modi analyse toegepast om de modi te identificeren die het karakter hebben van baseparen-opening. De geörienteerde DNA films voor de experimenten zijn gemaakt met een wet spinning methode en de instrumentatie is glansvol geinstalleerd en gemoderniseerd tijdens de promotie op het ILL.

APPENDICES

Appendix A

Theory and methods

A/1/ Modelling interactions in the solid state

To calculate physical or chemical properties of the solid state matter, such as total energy, equilibrium structure and dynamics of a set of atoms, several methods have been developed, some of which we describe below.

A/1/1/ Calculating energy of a system of atoms

A/1/1/1/ Empirical methods

These methods describe a biological system as a model of coupled oscillators where stiffness is governed by empirical inter-atomic potentials. These potentials are usually described by a list of force fields which take into account three energetic contributions:

- An intra-molecular part given by the sum of the bonding, angle and torsion energies of the oscillators.
- Two non-bonding parts given by the sum of the Coulomb and the Lennard Jones energies.

However in principle, these methods are not able to model chemical reactions and bond forming or breaking since the electronic structure of the system does not enter these models. Furthermore, there are still difficulties to treat hydrogen bond because this "strange bond" is governed by dispersion, electrostatic and covalent contributions and the increase of covalence between the hydrogen and the acceptor atoms, characteristic to certain short strong hydrogen bond, can give novel properties as proton transfer with temperature. We therefore used the ab-initio density functional theory method to calculate properties of our biological models.

A/1/1/2/ Ab-initio methods

For treating real biological systems, ab-initio methods use the approximation of Born-Hoppenheimer where electrons move around static nuclei. The purpose of all common

ab-initio (except DFT) methods is to solve the time independent Schrödinger equation $\hat{H}e\Psi = E\Psi$ by finding the total wavefunction of the electronic ground state $\Psi(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$ where $\{ \vec{r}_1, \vec{r}_2, ..., \vec{r}_N \}$ represents an ensemble of N electrons located at the $\vec{r}_1, \vec{r}_2, ..., \vec{r}_N$ positions respectively.

We will see how methods based on the Density Functional Theory find ground state energy by reducing the problem of knowing the many-body electronic wave function to a simpler problem of knowing the single-body electronic density [72].

The foundation of the theory of electronic structure of matter is the nonrelativistic Schrödinger equation for the many-electron wave function Ψ,

$$\left\{ -\frac{\hbar^2}{2m} \sum_j \nabla_j^2 - \sum_{j,l} \frac{Z_l e^2}{|r_j - R_l|} + \frac{1}{2} \sum_{j \neq j'} \frac{e^2}{|r_j - r_j'|} - E \right\} \Psi = 0$$

Where r_j are the positions of the electrons and R_l, Z_l the positions and atomic numbers of the nuclei.

A/1/1/2/1 Hartee-Fock method

This method treats electron as fermion particle, with the particularity that interchanging the electron label in the wave function $\Psi(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$ for the system of electrons has to change its sign. It means that the total wavefunction for the system is not a simple product of orbitals, but an antisymmetrized sum of all the products which can be obtained by interchanging electron labels. It is conveniently represented as a determinant (called Slater determinant):

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\vec{r}_1)\varphi_2(\vec{r}_1)......\varphi_N(\vec{r}_1) \\ \varphi_1(\vec{r}_2)\varphi_2(\vec{r}_2)......\varphi_N(\vec{r}_2) \\ .... \\ \varphi_1(\vec{r}_N)\varphi_2(\vec{r}_N)...\varphi_N(\vec{r}_N) \end{vmatrix}$$

Where $\varphi_i(\vec{r}_i)$ is the ith wavefunction describing the ith electron. So, Slater determinant satisfies the Pauli exclusion principle that each electron has to be described by a different wave function. The inclusion of the exchange interaction by means of the Slater determinant introduces exchange integrals in the calculation of electron-electron interaction where Coulomb interactions take usually part. The variational

principle stating that the wavefunction with the lowest energy is also the exact ground state wavefunction, is applied to find the lowest energy state of the system. The Schrödinger equation has now been reduced to a set of one-electron equations which can be solved numerically for a discrete set of positions or using a finite set of basis functions, such as Slater or Gaussian orbitals.

However, there are important reasons to not choose this method to calculate energies:

The Slater determinant which depend of the N electrons is complicated, and time for solving the Schrödinger equation grows as the power four of the number of electrons and therefore limit the number of atoms to be treated (around 10-20 for biological systems). The second reason is conceptually: in Hartree-Fock methods, each electron is treated as an independent particle moving around his nucleus in an effective average field created by the other. In reality, electrons correlate their motions trying to avoid each other. This interaction of correlation must be taking into account to avoid unphysical effect.

Now, there are many sophisticated methods for solving the many-body Schrödinger equation, e.g. there is diagrammatic perturbation theory in physics, while in quantum chemistry one often uses configuration interaction (CI) methods, based on the systematic expansion of the wavefunction in Slater multi-determinants. However, the problem with these methods is the huge computational effort (a full CI on N electrons is proportional to N!), which makes it virtually impossible to apply them efficiently to larger complex systems.

An alternative method to find energies of the electronic ground state consists of solving the Schrödinger equation by using the electron density which depends only of the variable **r**, instead of using the many-body wavefunction which depends of the 3N variables (or 4N if you count in spin).

To summarize, in Hartree-Fock methods, energy is a functional of the many-body wavefunction => E [$\Psi(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$] and in DFT methods, energy is a functional of the single-body density => E [$\rho(\vec{r})$].

A/1/1/2/2 Density Functional Theory

In DFT, the key variable is the electronic density $\rho(\vec{r})$ which is given by

$$\rho(r) = N \int d^3\vec{r}_2 \int d^3\vec{r}_3 .... \int d^3\vec{r}_N \Psi^*(\vec{r}, \vec{r}_2, ..., \vec{r}_N) \Psi(\vec{r}, \vec{r}_2, ..., \vec{r}_N)$$

The fact that the ground state properties are functionnals of the electron density $\rho(\vec{r})$ was proved by Hohenberg and Kohn in two fundamental theorems [73]:

    I.   Every observable of a stationary quantum mechanical system (including energy), can be calculated, in principle exactly, from the ground state density alone, i.e., every observable can be written as a functional of the ground state density.

    II.   The ground state density can be calculated, in principle exactly, using the variational method involving only density.

So, the ground state energy can be expressed:

$$E_0 = E[\rho_0] = \left\langle \Psi_0[\rho_0] \,|\, T + V + U \,|\, \Psi_0[\rho_0] \right\rangle$$

where the external potential created by nuclei $\left\langle \Psi_0[\rho_0] \,|\, V \,|\, \Psi_0[\rho_0] \right\rangle$ can be written explicitly in terms of the density $V[\rho] = \int V(\vec{r})\rho(\vec{r})d^3r$

The functionals T[ρ] and U[ρ] are called universal functionals while V[ρ] is obviously non-universal, as it depends on the system under study. Having specified a system, i.e. V is known, one as to minimise the functional $E[\rho] = T[\rho] + U[\rho] + \int V(\vec{r})\rho(\vec{r})d^3r$ with respect to $\rho(\vec{r})$, assuming one has got reliable expressions for T[ρ] and U[ρ]. A successful minimisation will yield the ground state density ρ_0 and thus all ground state observables as energy E_0.

The variational problem of minimising the energy functional E[ρ] can be solved by applying the Lagrangian method of undetermined multipliers, which was done by Kohn and Sham. Hereby, one uses the fact that the functional in the equation above can be written as a fictious density functional of a non-interacting system:

$$E[\rho] = T_s[\rho] + \int V_s(\vec{r})\rho(\vec{r})d\vec{r}\,, \text{ where } T_s \text{ denotes the non-interacting kinetic energy}$$

and V_s is an external effective potential in which the particles are moving. Obviously, $\rho_s(\vec{r}) \equiv \rho(\vec{r})$ if V_s is chosen to be $V_s = V + U + (T - T_s)$. Thus, one can solve the so-called Kohn-Sham equations of this auxiliary non-interacting system [74]

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_s(\vec{r}) \right] \phi_i(\vec{r}) = \varepsilon_i(\vec{r})$$

which yields the orbitals that ϕ_i reproduce the density $\rho(\vec{r})$ of the original many-body system $\rho_s(\vec{r}) \equiv \rho(\vec{r})$.

The effective single-particle potential V_s can be written in more details as

$$V_s = V + \int \frac{e^2 \rho_s(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3 r' + V_{XC}[\rho_s(\vec{r})]$$ where the first term is the nuclear potential, the

second term denotes the so-called Hartree term describing the electron-electron repulsion, while the last term V_{XC} is called exchange-correlation potential. Since the Hartree term and V_{XC} depend on $\rho(\vec{r})$ which depend on the ϕ_i, which in turn depend on V_s, the problem of solving the Kohn-Sham equations has to be done in a self consistent way. Usually one starts with an initial guess for $\rho(\vec{r})$, then one calculates the corresponding V_s and solves the Kohn-Sham equations for the ϕ_i. From these one calculates a new density and starts again. This procedure is then repeated until convergence is reached.

Approximations

The major problem with DFT is that the exact functionals for exchange and correlation are not known. First implementations of the Kohn-Sham method were using the local density approximation (LDA) to calculate the exchange correlation energy. This approximation is based upon exact exchange energy and from fits to the correlation energy for a uniform electron gas.

For historical reasons, the exchange correlation energy was divided into two parts:

$E_{XC}[\rho] = E_X[\rho] + E_C[\rho]$, the exchange energy which take into account the anti-symmetrisation of the wavefunction by exchanging electron and the correlation energy which is not known analytically, but is constantly improved on the basis of quantum Monte Carlo simulations, and fitted to analytical expansion.

In LDA methods, the functional depends only on the density at the coordinate where the functional is evaluated. Generalized gradient approximations (GGA) are still local but also take into account the gradient of the density at the same coordinate. This last method is the most appropriate to calculate physical properties of hydrogen bond system, since the electron density are more delocalized than other typical bond.

Furthermore long-range interactions as Van der Walls interactions are not calculated.

Advantages of this method

Traditionally ab-initio calculations have been performed on isolated molecules or small clusters containing several tens of atoms, but a number of computational codes, based on Density Functional Theory (DFT), now allow solids to be treated with an accuracy and computational cost proportional to N^3 comparable to that of the single

molecule calculations. The electronic correlation effect is not analytically known but is fitted to experimental data and take into account in the DFT calculations.

Solving Kohn-Sham equations

Several methods exists to solve these equations using planes waves, projector augmented waves, gaussian functions, molecular orbitals or other functions as basis sets. Plane waves such as used by the VASP package are the natural choice for periodic boundary conditions that are found in solid state. However, the cusp in the electron density at the position of the nuclei are difficult to describe by plane waves as many of them are needed, which can be overcome by the use of adjusted nuclear pseudopotentials or projector-augmented wave (PAW) potentials.

Using Bloch's theorem, each Kohn-Sham wavefunction can be written as the product of a plane wave and a cell-periodic part. Since the latter is a sum of plane waves, we get $\Psi_i(\vec{r}) = e^{i\vec{k}.\vec{r}} \sum_{\vec{G}} c_{i,\vec{G}} e^{i\vec{G}.\vec{r}} = \sum_{\vec{G}} c_{i,\vec{k}+\vec{G}} e^{i(\vec{k}+\vec{G}).\vec{r}}$

where **G** is the set of reciprocal lattice vectors, $c_{i,\vec{k}+\vec{G}}$ are the plane wave coefficients and **k** is the wavevector of the Bloch plane wave.

Electronic parameters we should have to worry

-Number of **k**-points:

In a periodic calculation, the number and positions in the reciprocal unit cell of the **k**-points is important for the accuracy of the calculation. To reproduce accurate dispersion of electronic energy, the common way is to get denser **k**-points meshes.

-Electronic density of states:

The number of **k**-points can be reduced by using a smearing method to take the partial occupancies of the states into account to calculate accurate band-structure energy.

In the "VASP-language", the flag ISMEAR determines how the partial occupancies $f_{n,\vec{k}}$ are set for each wavefunction to calculate charge density:

$\rho(\vec{r}) = \sum_{n} f_{n,\vec{k}} \left| \Psi_{n,\vec{k}}(\vec{r}) \right|^2$ and SIGMA determines the width of the smearing in eV. For non-insulator compound as for biological compound, a Gaussian smearing is more appropriated (ISMEAR=0). A width of 0.1 eV should be used for sensitive hydrogen bond systems because a too large width might result in a wrong total energy. However, small smearing parameters require a large k-point mesh.

-Cut-off energy for plane wave basis set:

All plane-waves with a kinetic energy smaller than a defined energy E_{cut} are included in the basis set, i.e., $E_{cut} = \frac{\hbar^2}{2m} G_{cut}^2$. Normally, for organic system, E_{cut} of hydrogen= 200eV. For more accurate calculations, this energy should be two times greater and it was experienced that transfer of proton is really simulated for an energy cut-off of 700eV as for pyridine-dicarboxylic acid (PDA) in a molecular dynamics simulation at room temperature.

A/1/2/ Finding equilibrium structure (T=0 K) by minimising forces acting on ions

An equilibrated structure is found at T=0 K and all the forces acting on ions are in principle 0. In ab-initio calculations, forces acting on ion (i) can be evaluated by knowing the variation of electronic Hamiltonian with the nuclei position and the ground state electronic wavefunction: $\vec{F}_i = -\langle \Psi_0 | \vec{\nabla}_i He | \Psi_0 \rangle$. These forces are the so-called Hellman-Feynman forces. It exist several methods to minimise these forces as the steepest descent or the conjugate gradient method.

A/1/3/ Probing the potential energy surface (PES) by moving ions:

A/1/3/1/ Normal modes methods

Normal modes of vibration are the eigenvalues and eigenvectors determined by diagonalising the dynamical matrix for each point in the reciprocal space situated in the first Brillouin zone.

Considering two elastic particles i,j with the mass m_i and m_j, the force due to j and acting on i is (with α,β one of the three spatial coordinates, $u_{i\alpha}$ the displacement of i along the bond):

$$\vec{F}_{i\alpha} = -\sum_j \sum_\beta \Phi_{\alpha\beta}(i,j).\vec{u}_{j\beta} = m_i \frac{d^2\vec{u}_{i\alpha}}{dt^2} = -m_i w^2 u_{i\alpha}(t) \quad (Newton)$$

as $\vec{u}_{i\alpha} = \vec{u}_{i\alpha 0} e^{-iwt}$

The force constants between these two particles can be found by applying the harmonic approximation of the potential energy $\Phi_{\alpha\beta}(ij) = -\frac{\partial Fi\alpha}{\partial u_{j\beta}} = \frac{\partial^2 V}{\partial u_{i\alpha}\partial u_{j\beta}}$. The

dynamical matrix can be constructed: $D_{\alpha\beta}(\vec{k}) = \dfrac{1}{m}\sum_{i,j}\Phi_{\alpha\beta}(ij).e^{i\vec{k}.\vec{R}_{ij}}$. It exists only one

solution if its determinant is null, i.e., $\sum_{\beta}[D_{\alpha\beta}(\vec{k}) - w^2\delta_{\alpha\beta}]A_{\beta} = 0$ (1). By

diagonalising this matrix for each point in the reciprocal space situated in the first

Brillouin zone, we obtain eigenvalues w^2, thus normal modes frequency w function of

wave vector **k** and mode j.

The phonon density of states tells us how many oscillators vibrate at the frequency w:

$$g(w) = \frac{1}{nd\Delta w}\sum_{k,j}\delta_{\Delta w}(w - w(\vec{k},j)) \quad with \ \delta_{\Delta w}(x) = \begin{cases} 1 \ if \ -\Delta\omega/2 \le x \le \Delta\omega/2 \\ 0 \ elsewhere \end{cases}$$

The eigenvalues and eigenvectors in equation (1) are determined by diagonalising the

dynamical matrix. Traditionally for molecular systems the dynamical matrix has only

been considered for the Gamma point, the Brillouin zone centre, giving calculations

that are valid for undispersed modes. Recently the treatment of the dynamical matrix

has been extended to the whole of the Brillouin zone to reproduce dispersion due to

intermolecular interactions and hydrogen bonds

A/1/3/2/ Born-Hoppenheimer Molecular dynamics methods:

In these methods, atoms are displaced according to Newton's equations of motion

knowing initial velocity $\vec{v}_i(0)$, Hellman-Feynman forces acting on ion i at each time

step, giving acceleration $\vec{a}_i(t) = \dfrac{\vec{F}_i(t)}{m_i}$, velocities $\vec{v}_i(t)$ and atomic position $\vec{r}_i(t)$ by

integration over time.

$$\vec{r}_i(t + \Delta t) = \vec{r}(t) + \vec{v}_i(0).\Delta t + \frac{1}{2}\vec{a}.\Delta t^2$$

$$\vec{v}_i(t) = \vec{v}_i(0) + \vec{a}.\Delta t$$

Initial random assignment of $\vec{v}_i(0)$ is given by the Maxwell distribution of velocities:

$\dfrac{1}{2}\sum m_i\vec{v}_i^2 = \dfrac{3}{2}Nk_BT$ where N is the number of atoms, k_B is the Boltzmann's constant

and T is the temperature in Kelvin. The integration time is a constant Δt=1fs, and

allows to measure the highest (Nyquist) frequency motion of $1/(2\Delta t)$ = 500 THz =

2.0677eV = 16 678cm^{-1}. The resolution in frequency is given by the inverse of the

total simulation time $\Sigma\Delta t$ multiplied by two: a simulation of 1ps gives a resolution of $0.5\text{THz} = 2.0677\text{meV} = 16.678\text{cm}^{-1}$.

The velocity auto-correlation function (VACF) is a memory function and is expressed as $C_{vv}(t) = \frac{1}{N}\sum_{\alpha=1}^{N} w_\alpha \frac{1}{3}\langle v_\alpha(0).v_\alpha(t)\rangle$ and the time Fourier transform gives the

vibrational density of states $G(w) = \int_0^{+\infty} dt \exp[-iwt]C_{vv}(t)$. This quantity can be directly compared to the phonon density of states.

In a DFT Born-Hoppenheimer molecular dynamics (especially in the VASP package), the electronic ground state is self-consistently reached at each time step giving Hellman-Feynman forces acting on ion i.

In MD simulation, some constant parameters must be set to define particular ensemble as microcanonical (NVE) or canonical (NVT) ensemble. In a NVE ensemble, the number of particles N, the volume of the simulation box and the total free energy (kinetic + potential) are constant.

In NVT ensemble, temperature is constant, i.e., average temperature over the simulation time must be equal to the initial guess. To perform it, a Nosé thermostat scales ion velocities at each time step and the induced temperature fluctuation show approximately the same frequencies as the typical phonon frequencies for the specific system. The period of this thermostatic bath could be critical since its corresponding frequency is in the low frequency range and could interfere with the modes of interest. For example, the thermostat's period (defined by the flag SMASS=0.1 in the INCAR file of VASP) can be around $75\text{fs} = 445\text{cm}^{-1}$ for hydrogen bond systems like urea-phosphoric acid and $61\text{fs} = 555\text{cm}^{-1}$ for PDA.

Complementarity of these methods

For the same total energy calculation, normal modes and molecular dynamics differ in the way of probing the potential energy surface (PES) and therefore to calculate vibrational frequencies of atoms. In normal modes methods, the potential energy surface is harmonically approximated. In MD methods, the real shape of PES is explored and is apparently the best choice to calculate vibrational frequencies of anharmonic potentials such as hydrogen bonds. However, normal modes methods has the real advantage to calculate the polarization eigenvector on atoms for each eigenvalues of vibrational energy and wavevector in the reciprocal space

A/1/4/ Typical VASP-input parameters (INCAR tag)

A/1/4/1/ Geometry optimisation

A/1/4/1/1/ INCAR.coarse (finding global minimum)

NELMIN = 4 (minimum of electronic self consistent cycles to reach the electronic ground state)
NELMDL = -3
EDIFF = 2E-4 (specifies the global break condition for the electronic SC-loop)
EDIFFG = -1E-2 (defines the break condition for the ionic relaxation loop)
IBRION = 2 (determines how the ions are updated and moved, a conjugate-gradient algorithm is used in this case)
LREAL =.TRUE. (determines whether the projection operators are evaluated in real-space or in reciprocal space, TRUE= projection done in real space)
NSW = 40 (number of optimisation steps)
ISMEAR= 0 (determines how the partial occupancies are set for each wavefunction)

A/1/4/1/2/ INCAR.fine (refine the global minimum)

$system = geom optimisation
NELMIN = 4
NELMDL = 0
EDIFF = 2E-6
EDIFFG = -1E-4
IBRION = 1 (a quasi-Newton algorithm is used to relax the ions into their instantaneous groundstate)
NSW = 50
ISMEAR= 0

A/1/4/2/ Molecular dynamics

A/1/4/2/1/ NVE ensemble

ALGO = F (convenient way to specify the electronic minimisation algorithm, in this case mixture of blocked Davidson block iteration scheme and RMM-DIIS algorithm)
IBRION = 0 (Born-Hoppenheimer molecular dynamics)
NELMIN = 4
NSW = 1000 (number of ionic steps)
POTIM = 1 (timestep in fs)
TEBEG = 50 (initial temperature in K)
ISYM = 0 (switch symmetry stuff off)
SMASS = -3 (determines the microcanonical ensemble)
LREAL = A (projection done in real space)
LCHARG =.FALSE. (determines whether the charge densities are written)

ISMEAR = 0
PREC=LOW (energy cutoff of plan waves)

A/1/4/2/2/ NVT ensemble

The same as for NVE ensemble except for SMASS>=0, a canonical ensemble is simulated using the algorithm of Nosé.

A/2/ Neutron scattering

In this section we will see how particle's correlation in space (**R**,t) are measured by neutrons in the reciprocal space (**Q**,w). In a neutron experiment, we measure the fraction of scattered neutrons with an energy E' in the range E' and E'+dE', and in a solid angle dΩ. It is called the doubly differential scattering cross section $\left(\dfrac{d^2\sigma}{d\Omega dE'} \right)$ and is proportional to the dynamical structure factor $S(\vec{Q}, w)$ which is simply the Fourier transform of the van Hove correlation function. To keep a constant neutron flux, we must take into account the factor $\dfrac{k'}{k}$, thus $\left(\dfrac{d^2\sigma}{d\Omega dE'} \right) = \dfrac{k'}{k} N\, S(\vec{Q}, w)$. It can be separated in two parts:

A coherent part which represent correlation between pair of atoms at time t=0 and time t: $\left(\dfrac{d^2\sigma}{d\Omega dE'} \right)_{coh} = \dfrac{1}{2\pi\hbar} \dfrac{k'}{k} \dfrac{\sigma_c}{4\pi} \sum_{n,n'} \int_{-\infty}^{+\infty} \left\langle \exp(-i\vec{Q}.\vec{R}_n(0)) \exp(i\vec{Q}.\vec{R}_{n'}(t)) \right\rangle \exp(-i\omega t) dt$

and an incoherent part which represent correlation of each atom with itself at time t=0 and time t: $\left(\dfrac{d^2\sigma}{d\Omega dE'} \right)_{inc} = \dfrac{1}{2\pi\hbar} \dfrac{k'}{k} \dfrac{\sigma_i}{4\pi} \sum_{n} \int_{-\infty}^{+\infty} \left\langle \exp(-i\vec{Q}.\vec{R}_n(0)) \exp(i\vec{Q}.\vec{R}_n(t)) \right\rangle \exp(-i\omega t) dt$.

Each partial cross section is the sum of an elastic and inelastic part. The elastic part represents for coherent scattering interferences between pair of atoms, one at time t=0 ant the other at infinite time. For incoherent scattering, it represents interferences between one atom at time t=0 and itself at infinite time.

A/2/1/ Investigating the structure of materials

Structure is defined as the position of one atom with his neighbour at equilibrium. Elastic coherent scattering, called diffraction, is the technique which allows to find the structure of a system of atoms.

A/2/2/ Investigating dynamics of materials

- By studying the translational or rotational motion in a confined space by determining the quasi-elastic neutron scattering. This particular scattering gives a broadening of the elastic peak and the width determines the energy of the characteristic motion.

The elastic incoherent scattering, proportional to the elastic incoherent structure factor EISF, is the correlation of one atom at time t=0 with itself at infinite time and give the possibility to see the geometry of its motion.

- By studying vibrations of atoms, the first approach is to determine the incoherent inelastic neutron scattering and the energy transfer is in the range (0-100meV)

In an inelastic neutron scattering (INS) experiment, information about molecular vibrations and lattice phonons is obtained from both energy and momentum transfer. The incoherent scattered intensity, relevant to hydrogenated matter, arising from a normal mode of vibration, j, for a system of n atoms, is proportional to the dynamic structure factor, expressed as for phonon calculations:

$$S_j(Q,\omega) \propto \frac{k_f}{k_0} \sum_{i=1}^{n} \frac{\sigma_i \hbar \left| \overline{Q}.\overline{C_{j,i}} \right|^2}{2 m_i \omega_k} \exp\left(-2 W_i^{'}(\overline{Q})\right) \delta\left(h\omega - h\omega_j\right) \qquad (1)$$

where ω_j is the mode frequency, $Q = k_f - k_i$ is the momentum transfer, $C_{j,i}$ is the eigenvector of atom i in mode j, m_i and σ_i are respectively the mass and scattering cross section of the atom i. $exp(-2W_i(Q))$ is the Debye-Waller factor of atom i which can be expressed $W_i = \frac{1}{2}(2\pi\overline{Q}) B_i (2\pi\overline{Q})$ where B_i is a symmetric 3x3 matrix representing the static correlation function of the displacements of atom i from equilibrium, that is the mean square displacement of the atom. Equation (1) shows that INS intensities can be directly calculated from a set of normal modes, therefore allowing a direct comparison between measured and calculated spectral profiles.

Spectral intensity is dominated by the contribution of hydrogen due to its large scattering cross-section (σ_h) and small mass, selective deuteration ($\sigma_d \sim \sigma_h/20$) therefore being a powerful tool for reducing scattering from particular molecular groups.

A/3/ Inelastic scattering instruments

A/3/1/ TOSCA

TOSCA [75] at ISIS is an inverted time-of-flight spectrometer in which the pulse of neutrons created by the spallation process covers the range of incident energies, which are determined by the time taken to reach the sample, and the final energy is determined by graphite analyser crystals. Spectra were measured from 20 to 4000 cm$^-$^{1}, with an energy resolution of ~1% of the energy transfer.

A/3/2/ IN1

IN1 [76] at ILL is a two-axis spectrometer in which the incident energy is varied as a function of the monochromator and its orientation, while the final energy is constant (determined by the beryllium filter). INS spectra were measured from 200 to 1400 cm$^-$^{1}, the resolution was about 5% of the energy transfer.

A/3/3/ IN4

IN4 [77] is a time-of-flight spectrometer used for the study of excitations in condensed matter. It works in the thermal neutron energy range (10-100 meV) and the resolution is between 2 and 5 % of the energy transfer

A/3/4/ IN5

IN5 [78] is a direct geometry time-of-flight spectrometer built on a cold-neutron source. The chopper system was set to yield an energy transfer range of $0.1 < E < 250$ meV, a momentum transfer range of $0.3 < Q < 1$ Å-1 and a resolution of 1%.

A/4/ Properties of hydrogen bonds

A/4/1/ Historic

In chemistry, a hydrogen bond is a type of attractive inter (or intra) molecular force that exists between two partial electric charges of opposite polarity. The nature of this kind of attraction is complex and still subject of discusses but it is now well establish that this particular bond play a central role in many aspects of molecular science.

A/4/2/ Physical and chemical properties of these particular bonds

This bond is formed by 3 atoms, X-H…X', XX' must be electronegative atoms (like N,O,F, Cl,….). X is the donor and X' the acceptor. The acceptor must have lone pair of electrons.

-interactions between X-H and X':

Electrostatic interactions between charged molecules (ex. H-F$^-$… H-F$^-$) ~ $1/r$

Dipole-dipole (Keesom) interactions via polarized bond (ex. H2O…H2O) ~ $1/r^3$.

Long range dipersive interactions type VDW in water cluster ~ $1/r^6$.

Possible covalent interaction between hydrogen and acceptor atom via lone pair of electrons of acceptor atom (ex. UPA crystal)

	normal	short	very short
bond energy (kJ/mol)	~1-2	~20	~40
donnor-acceptor length (Å)	2.8-3.0	2.6-2.8	2.4-2.6
stretching mode (cm-1)	~3000	~2200	~1500

Table 4

Thermodynamical, geometrical and vibrational properties of different types of HB

Appendix B

How Phonons Govern The Behaviour Of Short Strong Hydrogen Bonds in Urea-Phosphoric Acid

F. Fontaine-Vive, M.R. Johnson, G.J. Kearley, J.A.K. Howard and S.F. Parker, J.Am.Chem.Soc. 128, 2963 (2006)

B/1/ Abstract

Recent neutron diffraction data have shown that the hydrogen atom involved in the short, strong hydrogen bond in urea-phosphoric acid migrates towards the mid-point of the hydrogen bond as the temperature increases. With the help of solid state, ab-initio calculations and inelastic neutron scattering, we have investigated the temperature dependence of the structural and vibrational properties of the system. The potential energy surface of the proton in the short, strong hydrogen bond, and the thermal population of the energy levels therein, cannot account for the observed proton migration. Ab-initio, molecular dynamics simulations clearly reveal the migration of the proton. This molecular dynamics result was reported recently by other authors, but they only offered a tentative explanation in terms of a resonance between high frequency vibrations, which is not supported by the calculations presented here. We explain the proton migration in terms of phonon driven structural fluctuations and their impact on the temperature dependent evolution of the potential energy surface of the short hydrogen-bond proton.

B/2/ Introduction

Hydrogen bonds exhibit a wide range of structural and dynamic characteristics and play a fundamental role in many aspects of molecular science. These characteristics can be understood at a basic level in terms of a one-dimensional potential energy surface (PES). For a long hydrogen bond (donor-acceptor distance bigger than ~2.6Å), the PES is composed of two approximately harmonic potential wells, which are equally deep if the proton is equally stable in the equilibrium positions close to the donor and acceptor atoms. In the solid state, intermolecular interactions normally break this symmetry. As the hydrogen bond becomes shorter, the potential energy barrier separating them becomes lower and narrower until, for the shortest bonds

(donor-acceptor distance ~2.45Å), the potential wells merge and the single well is strongly anharmonic. Whereas long hydrogen bonds lead to weak perturbations of molecular properties, short, strong hydrogen bonds (SSHB) can have quite different properties. SSHB's are thought to play a fundamental role in stabilizing the intermediate state of certain enzymatic reactions through a subtle modification of the local environment of the hydrogen bond [12,13]. Other implications of SSHB's in inhibitor potency of active site[79] or in stabilization of the photoactive yellow protein [80] have been shown. DNA is an example of a hydrogen-bonded molecular system which contains the genetic code of living beings. Coding depends on the base-pair sequences, but also minor chemical changes to the nucleosides like the methylation of cytosine, which give rise to epigenetic effects[81]. In this context, the temperature dependent dynamics of protons in DNA hydrogen bonds and the chemical integrity of the base-pair molecules are vital to the reliability of genetic coding[82].

Measuring hydrogen bond specific properties in large biomolecular systems is very difficult and the level of quantum chemistry techniques would have to be reduced significantly when investigating such large molecular structures. Model molecular compounds are therefore required in order to identify and characterise SSHB's. Urea phosphoric acid (UPA) is such a system which has been extensively measured by single crystal, neutron diffraction as a function of temperature by Wilson and co-workers[83,84]. There are many hydrogen bonds in crystalline UPA, but only the SSHB shows novel properties, namely the migration of the proton from a stable position at low temperature (~ 50K) close to the donor, urea oxygen atom to a centred position in the hydrogen bond at 335K. The proton migrates by ~0.05Å.

Recently the same authors have published the results of ab initio molecular dynamics simulations which reproduce the measured proton migration effect with increasing temperature [[85]]. Understanding the proton migration mechansim from the MD simulations is not simple and a resonance between high frequency vibrations (1600 and 3200 cm^{-1}) was tentatively proposed as being related to proton migration. The authors suggest performing diffraction measurements on the deuterated analogue of UPA to test whether the proton migration is restricted when the vibrational resonance no longer exists. This type of work has been carried out on pyridine-3,5-carboxylic acid [16]. Proton migration is observed for both protonated and deuterated

compounds, which tends to invalidate an explanation in terms of vibrational resonances.

Furthermore, the authors show the time dependence of the structural variation of the SSHB (Fig. 3 of ref [85]). The fastest proton motion occurring on a timescale of ~25 fs (1300 cm^{-1}), is seen at the lowest temperatures (50K) and corresponds to the stretching vibration of the O-H bond. At higher temperature, the bond length is seen to be modulated by slower dynamics, but these were not characterised accurately in the short MD simulations (400 fs). The time dependence of the bond length indicate that lower frequency dynamics are responsible for proton migration and these have to be identified in order to understand correctly the novel behaviour of the SSHB.

The reported MD simulations suggest that proton migration could be induced by fluctuations in the local molecular environment, although a causal link between structural fluctuations and proton migration has not been demonstrated. Lattice expansion is not responsible for proton migration since this is small for a hydrogen-bonded solid and the published MD simulations were performed at different temperatures in the same unit cell. An alternative explanation is in terms of the Boltzmann population of the ground state and a low-lying vibrational energy level in the anharmonic potential well of the SSHB. Similar effects have been observed in normal hydrogen bonds, in which the two distinct proton positions, corresponding to the minima of the two-well potential, have populations that depend on temperature and the energy difference between the two tautomeric structures [86]. The form of the PES is of direct relevance to the recent paper of Morrison and co-workers since they have used a phenomenological model to interpret the structural evolution in terms of a two-well potential at low temperature and a single well potential at high temperature. The vibrational levels and therefore indirectly the PES can be measured using spectroscopic techniques like Raman, IR or inelastic neutron scattering (INS), although reliably assigning the vibrational modes for a hydrogen-bonded molecular crystal is a challenge.

Understanding vibrational spectra has in recent years been helped by the growing availability of quantum chemistry codes (notably based on density functional theory (DFT) methods) which perform normal mode analysis, that is by imposing the

harmonic approximation. Although less common, solid state DFT codes have been used in this context, which allow intermolecular interactions, including hydrogen bonds, to be incorporated [3,4,5]. Molecular vibrations typically calculated at the centre of the Brillouin zone can then be extended over the whole Brillouin zone, including dispersion [6,7,101]. Once vibration frequencies and displacement vectors have been obtained from the quantum chemical calculations, the INS spectra can be easily calculated given the scattering cross-sections of the atoms and the geometry of the spectrometer. In contrast to IR and Raman techniques, INS allows direct comparison between calculated and experimental spectra and therefore a clear evaluation of the numerical model.

Vibrational spectra can also be determined from molecular dynamics simulations (MD), the vibrational density of states being the Fourier transform of the velocity autocorrelation function and the neutron scattering functions $S(Q,w)$ being the Fourier transform of the van Hove correlation functions [87]. Compared to normal mode analysis (phonon calculations), MD has the advantage of not imposing the harmonic approximation and allows finite temperature to be investigated via the average kinetic energy of the atoms. The disadvantage of MD is that partial densities of states do not give as much mode specific information as normal modes and that the zero point motion of atoms cannot be included in a computationally efficient way. Provided the same total energy calculation is used, phonon and MD methods will therefore give complementary information.

We have measured the vibrational spectrum of UPA using INS and performed full phonon calculations on an optimised structure of UPA, evaluating the anharmonicity of these normal modes. We have explored directly the anharmonicty of the PES of the SSHB by calculating the potential energy variation for large amplitude displacements of the proton along the hydrogen bond. MD has been validated by comparing $S(Q,w)$ determined from a low temperature simulation (50K) with the results of the phonon calculations and then extended to higher temperature to further probe anharmonicity and the temperature-induced structural variations. Experimental and numerical results are reported in the following sections.

B/3/ Structural Information

The structure of urea-phosphoric acid (1/1), $C(NH_2)_2OH.PO_2(OH)_2$, was reported at a number of temperatures in the range 150-335K by Wilson [83]. Single crystal, neutron diffraction was used allowing all atomic positions to be accurately determined. UPA crystallizes in an orthorhombic structure, in the Pbca space group (Z=8). One lower temperature structure (15K) along with a geometry optimisation using DFT calculations was also reported [84]. Figure 39 shows the SSHB in detail with atom labelling consistent with that of Wilson.
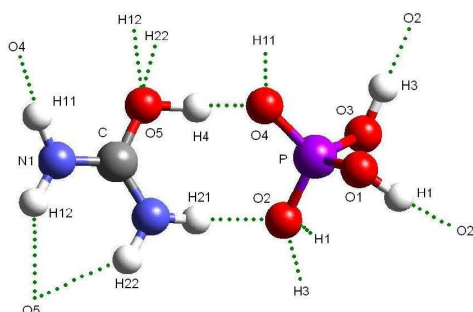

Figure 39

Urea and phosphoric acid molecules in their crystal configuration showing the SSHB O5-H4…O4.

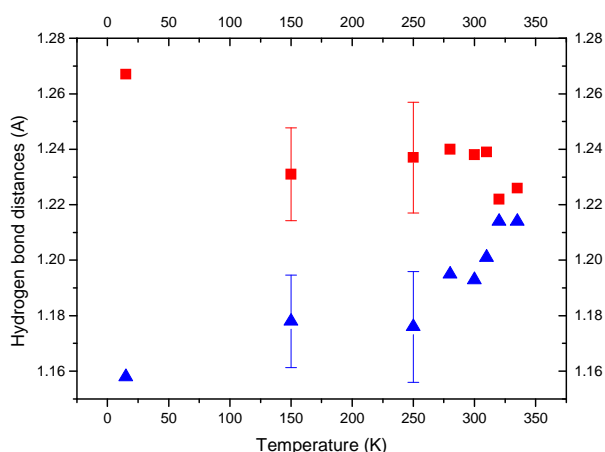The geometry of the SSHB O5-H4…O4 and the proton migration effect in the SSHB are shown in Figure 40.

Figure 40

Proton migration as a function of temperature in the SSHB of UPA as reported in[10].
Triangles represent the O5-H4 distance and the squares represent the O4-H4
distance. Typical error bars are shown at 150 K and 250 K.

As the basis for phonon calculations in this work, the measured structure of UPA has been optimised using the VASP DFT code [88,89,90,91]. Since the implementation of DFT in computational codes does not include long range dispersive interactions, optimising the atomic coordinates in the crystallographically determined unit cell can give a better description of the structure, and thereafter the dynamics of the system, than optimising the cell parameters. This is seen to be the case here when comparing measured and calculated INS spectra (below). The measured crystal structure is very close to a potential energy minimum in the VASP calculation, with bond geometries only being slightly modified. The biggest difference occurs for the SSHB, which has an O-O length of 2.43 Å an O-H length of 1.10 Å and an OHO angle of 168.2 degrees, compared to the corresponding experimental values of 2.40 Å, 1.16 Å and 170.0 degrees respectively. The DFT calculation tends to underestimate the strength of the hydrogen bond giving a shorter O-H distance and a longer O-O distance so that the offset of the proton position from the middle of the hydrogen bond is about twice as big as in the measured structure. When optimising the crystal structure in the unit cells determined experimentally at higher temperature, the most energetically stable structures (corresponding to the global minimum of the PES) retain a short O-H

distance, which precludes a simple explanation of proton migration in terms of lattice expansion.

B/4/ Vibrational spectroscopy, normal mode analysis and molecular dynamics simulations

Figure 41 (upper curve) shows the INS spectrum for the fully protonated sample of UPA, measured on TOSCA at the ISIS facility (UK) [75]. Often for a hydrogen-bonded system, it is straightfoward to prepare partially deuterated samples, which, due to the scattering cross-section of deuterium being about 20 times smaller than that of hydrogen, allow spectral intensity from specific atoms to be attenuated or highlighted. However, in UPA, all hydrogen atoms are labile and exchange with deuterium, so no useful contrast variation could be obtained. The figure also shows two calculated spectra; the bottom curve is from the phonon calculation [69] and the middle curve is from the MD simulation at 50K. Both calculations are in reasonably good agreement with the experimental data in terms of the accuracy of spectral frequencies and the overall intensity distribution. The main discrepancy concerns the intense mode at 800 cm^{-1} in the data, which is found at 850 cm^{-1} in the spectrum calculated from the MD simulation and 890 cm^{-1} in the spectrum from the phonon calculation. The normal mode at 890 cm^{-1} is seen to involve essentially a wagging motion of the OH groups in the acid molecule. The discrepancy of 40 cm^{-1} between the two calculations is due to anharmonicity and the remaining difference is due to the functional (PW91)[92] for the exchange-correlation energy and Vanderbilt-type ultrasoft pseudopotentials[93] for ion-electron interactions used in all calculations.
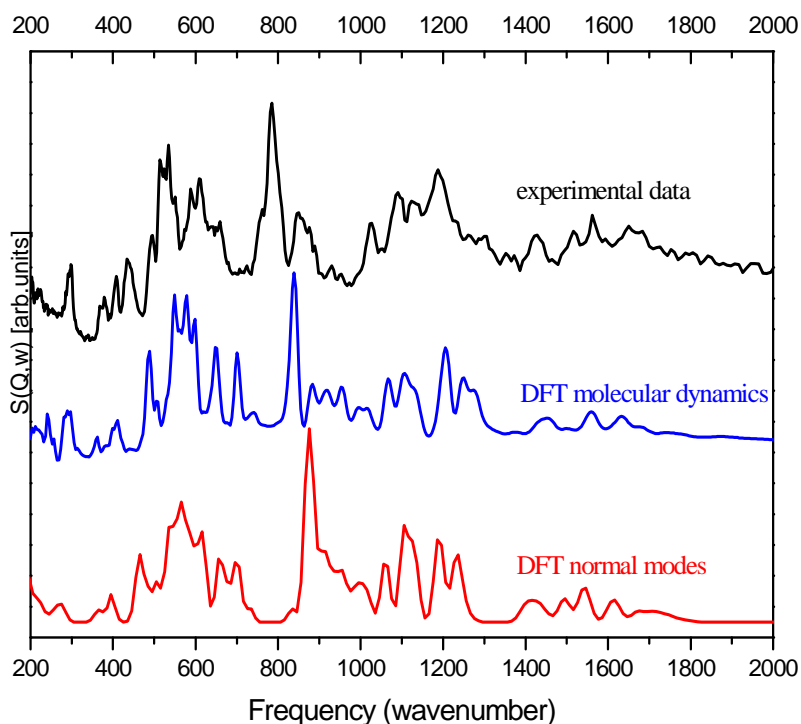
Figure 41

Comparison of measured (top) and calculated (middle: molecular dynamics and bottom: normal modes analysis) INS S(Q,w) spectra of urea-phosphoric acid.

The mode assignment at 800 cm^{-1} is confirmed by the partial, vibrational density of states, which shows the striking differences between the normal and short hydrogen bonds that could not be accessed experimentally with partially deuterated samples (Figure 42). The lower panel shows the partial density of states of the hydrogen atoms in O1-H1…O2 type bonds. The density of states extends mainly from ~400 cm^{-1} to 2800 cm^{-1}, the small downshift in the OH stretching frequency from above 3000 cm^{-1} being due to the weak transfer of electron density from the covalent bond to the hydrogen bond.

The upper panel shows the partial density of states for the SSHB hydrogen atom. The intensity in this case is contained in a much smaller frequency range, extending from 1100 cm^{-1} to 1800 cm^{-1}, with the bands at 1100 cm^{-1}, 1500 cm^{-1} and 1600 cm^{-1} involving wagging motion of the OH bond (which is strongly coupled to H-N-H bending vibrations) and the bands above 1700 cm^{-1} involving mainly stretching motion. That the lower frequency limit for the density of states is at a relatively high frequency is another indication of the significant electron transfer from the covalent to

the hydrogen bond [94,95]. These assignments are consistent with the analysis of infrared and Raman results [96]. At lower frequency, the vibrational character is more difficult to describe since the vibrations of the urea and phosphoric acid molecules are strongly coupled by the SSHB and the other hydrogen bonds.

As for S(Q,w) (Figure 41) there is reasonably good agreement between the densities of states obtained from the phonon calculation and MD simulation (lower two curves in each panel of Figure 42). From this comparison, anharmonicity is most apparent for the SSHB for the mode at 1400 cm^{-1}. However, the anharmonicity is not apparent in S(Q,w) because it includes the signal from all atoms, in particular from the urea NH_2 groups in the 1400 – 1500 cm^{-1} range. The temperature dependence of the densities of states does not reveal any further, significant anharmonic effects, especially for the normal hydrogen bonds.
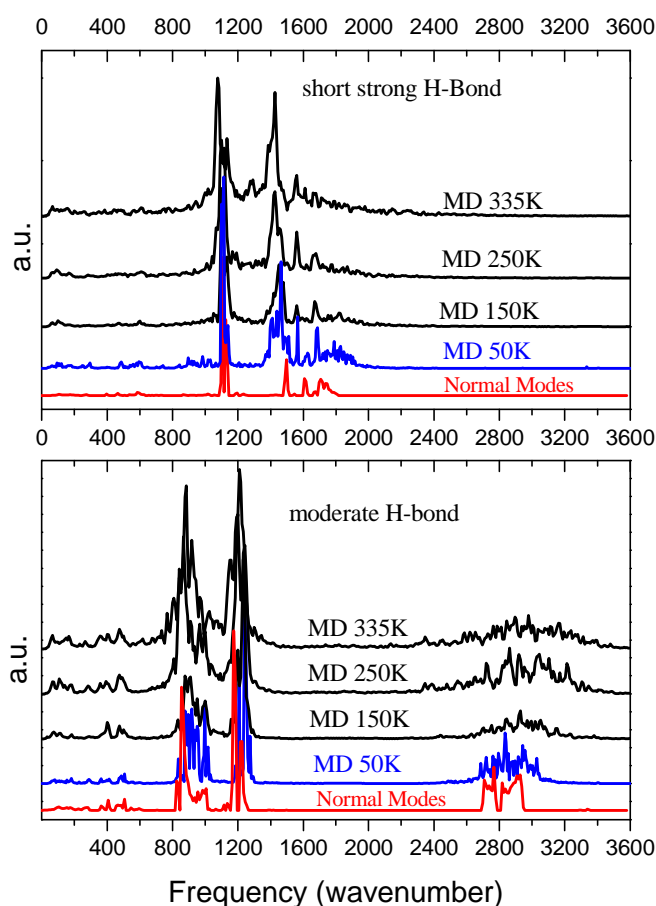

Figure 42

The vibrational density of states of the hydrogen involved in the strong H bond O5-H4...O4 (upper panel) and of the H involved in a moderate H bond O1-H1...O2

(lower panel). The lowest curve in each panel is derived from phonon calculations, the upper four curves are from MD simulations.

B/5/ Anharmonicity; Mapping the PES

Anharmonicity is a recurring theme in studies of systems with SSHB's [97]. In addition to the analysis in the preceding section, the anharmonicity of the normal modes can be investigated by calculating the PES for structures obtained by applying normal mode displacements to the equilibrium structure. The maximum amplitude of the displacements should induce an increase in potential energy equivalent to the corresponding vibrational frequency (for example ~2 times the vibrational energy). We have analysed every normal mode for UPA in the frequency range from 900 cm^{-1} to 2000 cm^{-1} in this way, fitting each 1D slice of the PES along the normal coordinate with a 4th order polynomial and using the fit parameters in a perturbative correction to the harmonic frequency. The biggest anharmonic correction was less than 1.2% of the harmonic frequency.

A common way to evaluate anharmonicity is to calculate eigen-energies and functions in the 1D slice of the PES of the O5-H4 stretching displacement along the SSHB, the potential energy variation being determined from the variation in the crystal energy as a function of the position of the H4 proton. From the low temperature structure, the calculated potential energy well is broad and asymmetric with a single minimum close to the O5 atom (Figure 43). It is not a two-well potential. The energy levels and wavefunctions in this potential energy well have been obtained from a numerical solution of Schroedinger's equation, assuming the oscillator mass to be that of a proton, giving a vibrational frequency of 1525 cm^{-1}, compared to ~1700 cm^{-1} in Figure 42 (upper panel). However, phonon calculations show clearly that the O5-H4 stretch is not a normal mode which has two consequences; the effective oscillator mass is not simply that of the proton and the spatial exploration of the anharmonic potential energy well by any normal mode is limited. Accordingly, a harmonic analysis of the potential energy well close to the minimum (± 0.05 Å) gives a vibrational frequency of 1555 cm^{-1}, compared to 1525 cm^{-1}.

Mapping a 1D slice of the PES as in Figure 43 does show that there cannot be excited vibrational energy levels associated with the SSHB that could be populated at room temperature and therefore explain the observed proton migration. While quantifying the vibrational energies in the PES depends on the oscillator mass, the partial density of states for the SSHB proton (Figure 42, upper panel) shows that these range between 1700 cm^{-1} and 1800 cm^{-1}, that is between 2400 K and 2500 K.
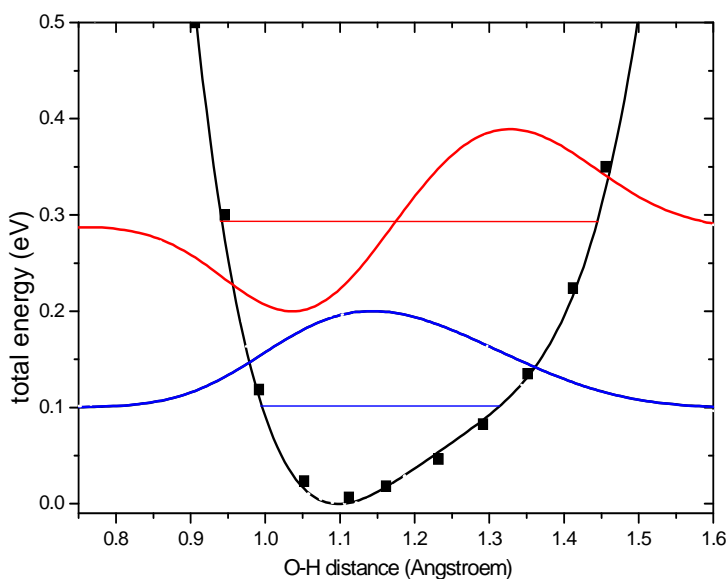

Figure 43

Potential energy well, vibrational energy levels and wavefunctions of the hydrogen atom in the SSHB. The wavefunctions represent the fundamental state n=0 (lower curve) and the first excited state n=1 (upper curve) in the anharmonic potential energy well, the vibrational energy is 1525 cm^{-1}.

B/6/ Structural analysis of UPA as a function of temperature

Having ruled out proton migration arising from simple lattice expansion or from the population of energy levels in the PES of the SSHB, the relevant structural information has to be sought in the MD simulations as a function of temperature. From the atomic trajectories a histogram of the O5-H4 distance can be constructed at each temperature. Figure 44 shows Gaussian fits to these histograms. The average O-H bond length increases from 1.1 Å to 1.18 Å in going from 50K to 335K and the full width of the distribution increases by a factor of ~4. This temperature dependent evolution is in reasonably good agreement with the experimental observations, the

corresponding values being 1.16 Å and 1.21 Å, and the recently published MD results. Since the O-O distance in the SSHB is slightly longer and the OH bond is slightly shorter in our calculations than in the measured structure, the proton does not quite reach the middle of the hydrogen bond, by 0.04 Å, at a temperature of 335K. Increasing the temperature in the MD simulation does not allow the proton to migrate further. The lack of zero point motion in MD simulations, which, according to phonon calculations, contributes ~25% of the mean square displacement of the H4 proton at 335K, is also responsible for the reduced amplitude of the calculated proton migration. So having eliminated an explanation in terms of the static PES of the hydrogen bond, MD reveals the mechanism of the proton migration to be fluctuations of the molecular environment, although the amplitude of the effect is underestimated.
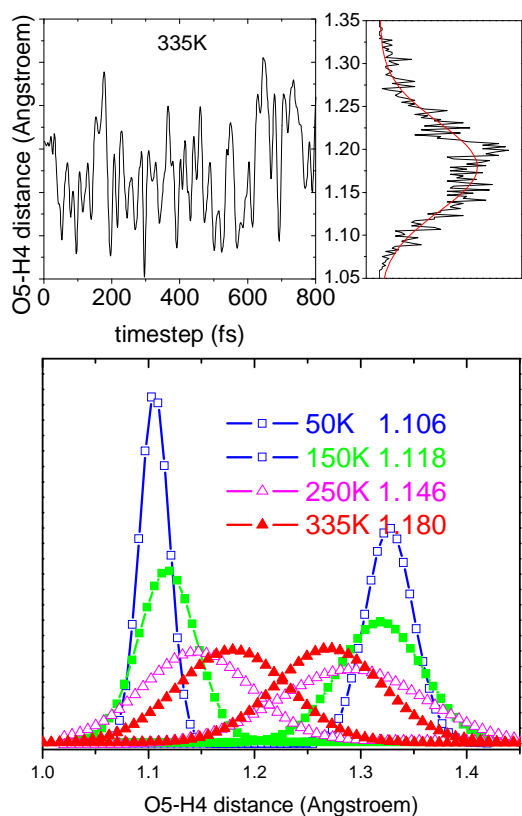

Figure 44

The upper panels show (left) the O5-H4 distance extracted from part (0.8 ps of a total of 2 ps) of the MD simulation at 335K and (right) a histogram representation of the whole trajectory and a Gaussian fit. The lower panel shows the Gaussian fits to the O5-H4 and the O4-H4 data from the MD simulations at four temperatures.

Additional understanding of the proton migration mechanism can be obtained by combining structures from MD simulations with methods for mapping the PES. By monitoring the distance C-P between the urea carbon atom and the acid phosphor atom during the MD simulations, crystal structures (symmetry P1) have been extracted which correspond to the maximum and the minimum of the C-P distance. For each of these structures the 1D slice of the PES, like the one shown in Figure 43, has been calculated (see Figure 45). At 150K, the two PES's are similar and ressemble the one determined from the equilibrium structure (Figure 43). At 335K there are bigger differences between the two PES's due to the thermal fluctuations and, more importantly, both PES are relatively flat-bottomed compared to the lower temperature PES's. This shape of the PES allows the proton to spend more time closer to the acid oxygen while, for the PES for the maximum C-P distance, the potential energy minimum is actually shifted by ~0.07 Å away from the urea molecule.

The high temperature structure, with the proton shifted towards the centre of the hydrogen bond, is stabilised by vibrational entropy, that is by the low frequency lattice vibrations that are activated up to 335 K. A time-independent potential energy well can be obtained from the time-averaged crystal structure and is similar to the average of the two, high temperature potential energy wells in Figure 45. The proton is shifted by ~0.06 Å away from the urea oxygen atom. At any temperature, fluctuations in the potential energy well tend to broaden the distribution of proton positions. Figure 45 shows directly the change in the average potential energy well and its time-dependent modulation for the SSHB due to thermal fluctuations and therefore the mechanism for proton migration.
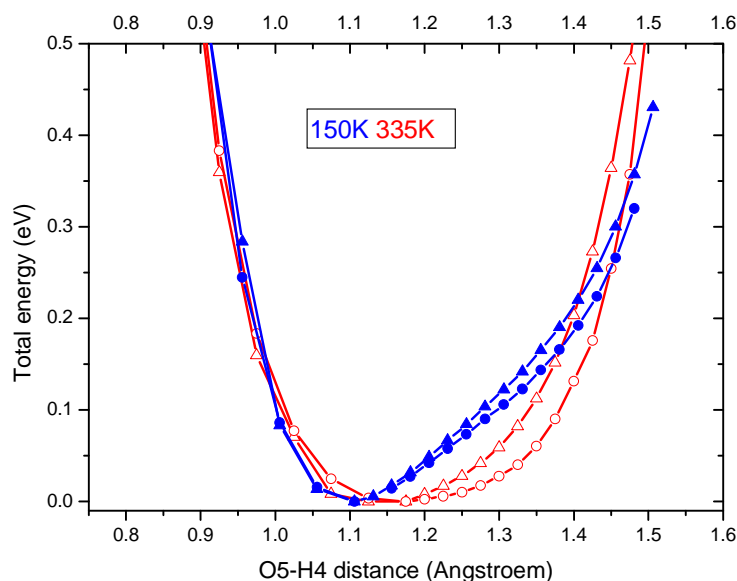
Figure 45

The potential energy wells for the O5-H4 stretching displacement at 150 K (filled symbols) and 335 K (open symbols) temperatures for the crystal structures with the minimum (triangles) and maximum (circles) separations of the urea and phosphoric acid molecules.

The final question concerns the characteristics of the dynamics that induce proton migration. Since this effect occurs continuously in the temperature range from 50K to 335K, the lattice vibrations (phonons) of the crystal are expected to be responsible. The frequencies that modulate the O5-H4 bond length can be found from the Fourier transform (FT) of the time dependence of the bond length variation. Figure 46 shows the amplitude of the FT which therefore displays the amplitude of the oscillations as a function of frequency. This figure clearly reveals that the lowest frequency modes ($\sim$50 - 200 cm^{-1}), that is the phonons, give rise to the strongest modulation of the O5-H4 bond length. We note that the amplitude of the FT tends to zero at zero frequency since the long wavelength acoustic phonons do not modulate the local structure.

Dynamic properties of hydrogen bonds can be influenced by specific vibrational modes. For example, in benzoic acid dimers, proton transfer is enhanced by a small number of modes and, in particular, the intra-dimer stretch. In UPA, upto 200 cm^{-1}, there are 80 normal modes. If any of these modes in particular were involved in the modulation of the potential energy well, a well-defined peak could be expected in the

FT's of Figure 46 (the MD trajectories are sufficiently long and the frequency resolution is good enough, see for example Figure 42). Figure 46 therefore suggests that many modes are involved in the modulation of the potential energy well. Inspection of the modes (eigenvectors) reveals a certain similarity in that they tend to involve molecular rotations which sollicit several hydrogen bonds, reflecting the complex 3D hydrogen bond network in UPA. This situation is more complicated than benzoic acid, which can be considered as an isolated dimer with only one type of hydrogen bond.
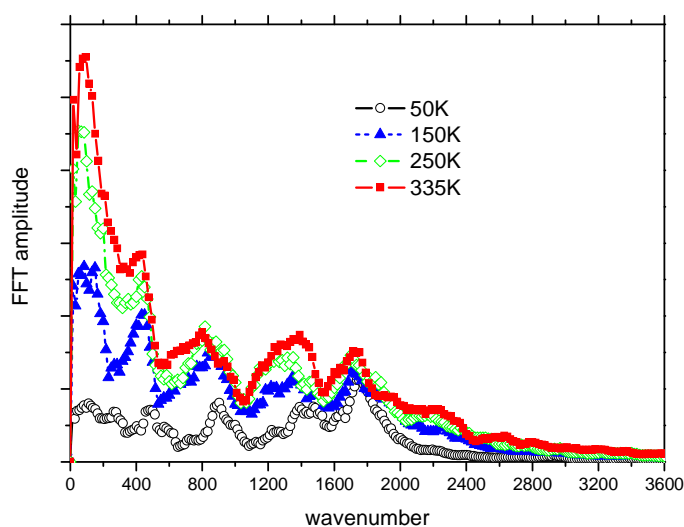

Figure 46

The FFT amplitude of the O5-H4 bond length variation at four temperatures.

B/7/ Discussion

Similar MD simulations were recently reported by Morrison and co-workers [85]. While their simulations were long enough to reproduce the proton migration with temperature, they were too short to obtain adequate resolution in the frequency spectra that they extracted. Their resolution was of the order of 100 cm^{-1} whereas we have improved this by a factor of ten or more. Furthermore they have taken the Fourier transform of the forces acting on the atoms whereas the physically relevant quantity is the vibrational density of states, which is obtained from the velocity auto-correlation function. Their tentative explanation of proton migration being driven by a resonance between one and two-phonon excitations of different vibrations is therefore highly speculative since, firstly, there are many possible such resonances and

secondly, the density of states for the modes they cite at ~1600 cm^{-1} is very weak (see Figure 42). Furthermore if these vibrations were responsible for the significant changes in the hydrogen bond geometry, jumps in the proton position would occur on a time scale shorter than 25 fs, corresponding to the approximate O5-H4 stretch frequency, whereas figure 2 of [85] and Figure 44 in this article show that these jumps occur on a longer time scale. These slower dynamics are quantified in the Fourier transform of the O5-H4 bond length and are seen to correspond mainly to the lattice vibrations in the 50-200 cm^{-1} frequency range.

The work presented in this paper also shows that the phenomenological model used in [85] to interpret the MD results is not appropriate in the stated form. The authors talk about a transition from a two-well potential to a single well potential, whereas we have clearly shown that the PES always has a single minimum and that proton migration is due to rather subtle change in the average PES induced by the fluctuating environment.

B/8/ Conclusion

With a view to understanding the proton migration in the SSHB in UPA we have calculated and measured, directly and indirectly, the PES that governs the proton dynamics. INS has been used to measure the vibrational modes of UPA at 20 K. Characteristics of the PES are obtained from solid state DFT calculations that are validated by comparing calculated spectra with the measured spectrum. Based on this comparison, the phonon calculation, based on the harmonic approximation, can be seen to give a good description of the vibrational modes in UPA. Further analysis reveals clear differences in terms of vibrational modes, between the SSHB and the other, normal hydrogen bonds. The normal hydrogen bonds have modes that cover a much larger spectral range, from 800 cm^{-1} for wagging modes to 2500 cm^{-1} for stretching modes. In contrast, the SSHB proton participates in modes that are concentrated between 1100 cm^{-1} and 1800 cm^{-1} revealing a significant transfer of electron density from the covalent bond to the hydrogen bond. Since there is generally a strong vibrational density of states in the spectral range up to ~2000 cm^{-1}, there is a strong mixing of the SSHB vibrations with other molecular vibrations.

The first and simplest step in the analysis of the PES (and in the determination of the phonons) of UPA is to find the global minimum, which corresponds to the low temperature crystal structure. Taking the measured unit cell at any temperature, or the unit cell optimised by DFT, gives a short O-H bond length of 1.1 Å. Proton transfer in UPA is not a simple consequence of lattice expansion.

The DFT calculations are also used to map out the 1D slice of the PES of the proton in the SSHB in the O5 – O4 direction. As expected, the potential energy well is strongly anharmonic, but this does not show up clearly in the vibrational spectroscopy because the mixed character of the normal modes means that only a limited spatial range of the PES is explored. Vibrational energy levels in the potential energy well have been calculated in the harmonic approximation and taking into account the anharmonicity. Although the excited state wavefunction in the anharmonic potential has the centre of gravity of the corresponding probability distribution shifted towards the centre of the hydrogen bond, the excited state is at least 2400 K in energy above the ground state and cannot be populated at room temperature.

Molecular dynamics simulations using the same DFT code, VASP, also give a vibrational spectrum that is in good agreement with the INS spectrum. MD is not restricted to the harmonic approximation and shows that anharmonicity does not significantly affect the vibrational modes in UPA. Analysing the time and temperature dependence of the molecular geometries from the MD trajectories shows clearly that the proton migration is due to thermal fluctuations of the molecular environment induced by the lattice vibrations. The high temperature structure is stabilised by vibrational entropy and the resulting time-averaged PES has a minimum that shifts towards the centre of the hydrogen bond and the shape becomes more "flat-bottomed", which also favours proton migration. Time-dependent fluctuations broaden the distribution of proton positions about the average value. MD simulations under-estimate slightly the extent of proton transfer, which is due to the under-estimated hydrogen bond strength and to the protons being treated as classical particles.

This combined neutron scattering and DFT modelling work demonstrates clearly that the properties of short hydrogen bonds can be modified by the thermal fluctuations of

the local molecular environment, that is by phonons in the case of a molecular crystal like UPA. Temperature, weak inter-atomic interactions and local structure collectively govern the reactivity of SSHB. Changing any of these parameters can therefore modify the properties of the SSHB and the processes that they mediate, like enzymatic reactions and inhibitor potency.

Appendix C

Phonon Driven Proton Transfer in Crystals with Short Strong Hydrogen Bonds

F. Fontaine-Vive, M.R. Johnson, G.J. Kearley, J. A. Cowan, J.A.K. Howard and S.F. Parker, the Journal of Chemical Physics, 124, 234503 (2006)

C/1/ Abstract

Recent work on understanding why protons migrate with increasing temperature in short, strong hydrogen bonds is extended here to three more organic, crystalline systems. Inelastic neutron scattering and density functional theory based simulations are used to investigate structure, vibrations and dynamics of these systems as a function of temperature. The mechanism determined in previous work on urea-phosphoric acid of low frequency vibrations stabilising average crystal structures, in which the potential energy well of the hydrogen bond has its minimum shifted towards the centre of the bond, is found to be valid here. The new feature of the N-H…O hydrogen bonds studied in this work is that the proton is transferred from the donor atom to the acceptor atom. Molecular dynamics simulations show that, in an intermediate temperature regime, in which the proton is not completely transferred, the proton is bi-stable, jumping from one side of the hydrogen bond to the other. In the case of 3,5-pyridine dicarboxylic acid, which has been studied in most detail, specific phonons are identified, which influence the potential energy surface of the proton in the short, strong hydrogen bond.

C/2/ Introduction

Hydrogen bonds exhibit a wide range of structural and dynamic characteristics and play a fundamental role in many aspects of molecular science. Whereas many hydrogen bonds can be thought of as giving rise to a modest perturbation of the properties of molecules, short, strong hydrogen bonds (SSHBs) constitute a much stronger perturbation, which can result in unique physical or chemical behaviour. SSHBs are thought to play a fundamental role in stabilizing the intermediate state of

certain enzymatic reactions through a subtle modification of the local environment of the hydrogen bond [12,13]. Other implications of SSHBs in inhibitor potency of active site [79] or in stabilization of the photoactive yellow protein [80] have been shown. DNA is an example of hydrogen bonded molecular system which contains the genetic code of living beings. Coding depends on the base-pair sequences, but also minor chemical changes to the nucleosides like the methylation of cytosine, which give rise to epigenetic effects [81]. In this context, the temperature dependent dynamics of protons in DNA hydrogen bonds and the chemical integrity of the base-pair molecules are vital to the reliability of genetic coding.

The unique property of SSHBs of interest in this paper is the proton migration with increasing temperature away from the donor atom towards or beyond the centre of the hydrogen bond. This effect has been observed by neutron diffraction in a number of compounds; urea phosporic acid (UPA) [83,84], pyridine-3,5-dicarboxylic acid (PDA) [15,16], 1:2 co-crystal of benzene-1,2,4,5-tetracarboxylic acid and 4,4'-bipyridyl (BTA-BPY) [98] and 1:1 crystal of 2-methylpyridine and pentachlorophenol (4-MePy-PCP) [99].

Temperature dependent proton migration can be thought of as being structurally or vibrationally driven. Lattice expansion could be sufficient to modify the geometry of the SSHB in an appropriate way or there could be a subtle structural reorganisation with temperature that favours a centred position of the proton in the hydrogen bond. Otherwise depending on the shape of the potential well of the hydrogen bond there could be excited vibrational states, which can be thermally populated, with a centre of gravity shifted towards the centre of the hydrogen bond.

Recent inelastic neutron scattering (INS) and ab initio numerical simulations on UPA [100] showed that vibrational levels in the potential well of the proton in the hydrogen bond could not be significantly populated at temperatures at which proton migration is observed. Total energy calculations based on density functional theory (DFT) methods showed the structures with a short donor-H distance to be most stable whatever the size of the unit cell. Molecular dynamics (MD) with DFT revealed the proton migration effect. Analysis of the simulation trajectories showed low frequency lattice vibrations to be responsible for stabilising the crystal structure at non-zero

temperatures in such a way that the minimum of the average potential shifts towards the centre of the hydrogen bond. The purpose of this paper is therefore to investigate whether the same mechanism drives proton migration in the other SSHBs.

UPA is a complicated system in that the crystal structure has many inter-molecular hydrogen bonds, which involve all hydrogen atoms, and only one of which is short. As a result of this 3D coupled network, particular phonons could not be identified as driving proton migration. The three other systems presented here have simpler hydrogen bond networks; 2D sheets in the case of PDA, 1D chains for BTA-BPY and dimers for 4MP-PCP. In contrast to UPA in which the SSHB is a O-H…O bond, the three samples here present short N-H…O bonds. In these three systems, there are fewer hydrogen bonds, which allow the hydrogen bond vibrations to be highlighted via selective deuteration in INS experiments. The three systems are also more pronounced examples of proton migration, the proton being transferred beyond the centre of the hydrogen bond from donor to acceptor atoms. In this case does the vibrationally-stabilised, time-averaged potential well take the proton continuously from donor to acceptor as a function of temperature?

In this paper we focus mainly on PDA for which we have used INS to investigate the molecular and hydrogen bond vibrations. These measurements are confronted with phonon calculations based on DFT methods, which allow the vibrational modes to be assigned, and mapping and anharmonic analysis of the hydrogen bond potential well. MD simulations using the same DFT methods allow temperature dependent effects to be investigated that are not constrained to the low temperature crystal structure via the harmonic approximation, like thermal amplitudes determined from phonon calculations. For the other two samples, BTA-BPY and 4MP-PCP, only MD simulations have been performed to investigate directly proton migration. Experimental and computational details are described in the next section and thereafter results are presented for each of the three systems.

C/3/ Experimental details

PDA was obtained from Sigma-Aldrich and used in protonated form without further purification. The hydrogen bonds were deuterated by multiple recrystallisations from D_2O. INS spectra were measured at a temperature of ~20 K on the inverted geometry spectrometers, IN1 at ILL [76] and TOSCA at ISIS [75].

C/4/ Numerical details

All DFT calculations were performed using the VASP DFT code [88,89,90,91]. We used PAW pseudopotentials with a plane wave cutoff of 700 eV in combination with the PBE functional [92]. This energy cutoff is much higher than we normally use (270 eV) and is found to give a significantly better description of the SSHBs. The k-point spacing was typically 0.1 Å^{-1}.

As for organic crystals [3,4,5,6,7] and biopolymers [101], phonon calculations were performed for PDA using the direct method as implemented in the PHONON code [69]. A supercell is constructed from the optimised unit cell structure, which allows all non-zero force constants to be determined when the crystallographically inequivalent atoms are displaced from their equilibrium positions. The force constants from VASP calculations are used to generate the dynamical matrix in PHONON, which is then diagonalised for a large number of randomly chosen points in the reciprocal lattice to obtain the vibrational density of states (vDOS). S(Q,w) is calculated from the vDOS in the incoherent approximation for the specific geometry and resolution of each spectrometer.

Molecular dynamics simulations were also performed with VASP. The results presented here were obtained in the NVE ensemble, that is without coupling to an external heat bath. A time step of 1 fs was used and the total simulation time was typically 5-10 ps. vDOS were extracted from MD trajectories using nMoldyn [87] and otherwise geometrical information was obtained directly from the time-dependent atomic coordinates.

The potential energy well for the proton in the hydrogen bond was determined for any given structure by calculating the crystal energy as a function of the displacement of the proton along the donor-acceptor direction. The energy levels within a potential

well were determined from a numerical solution of Schrödinger's equation assuming the mass of the oscillator to be that of the proton.

C/5/ Results

C/5/1/ Pyridine-3,5-dicarboxylic acid

C/5/1/1/ Structure

The structure of pyridine-3,5-dicarboxylic acid was reported at a number of temperatures in the range 15 K - 296 K by Cowan [15,16]. PDA crystallizes in the $P2_1/c$ space group with four molecules in the unit cell. One short strong N1-H5…O4 and a "normal" O-H…O hydrogen bond link the molecules in two-dimensional planar sheets in the *a-b* plane.

	N1….O4	N1….H5	H5….O4	N1-H5-O4
exp data				
296K	2.525	1.308	1.218	177.9
100K	2.526	1.181	1.361	172.3
15K	2.523	1.213	1.311	176.4
MD simulations				
300K	2.541	1.206	1.331	169.3
250K	2.546	1.162	1.371	170.1
15K	2.535	1.128	1.409	174.2
optimised geometries				
15K	2.551	1.122	1.429	176.5

Table 5

Proton transfer as a function of temperature in the N1...H5...O4 SSHB in PDA as reported in [15,16] and obtained from simulation.

	O1….O3	O1….H4	H4….O3	O1-H4-O3
exp data				
296K	2.591	1.004	1.608	165.1
15K	2.569	1.025	1.558	167.7
MD simulations				
300K	2.564	1.037	1.546	164.2
15K	2.516	1.049	1.471	172.1
optimised geometries				
15K	2.471	1.091	1.383	175.3

Table 6

O1...H4...O3 Hydrogen bond geometry as a function of temperature in PDA as reported in [15,16]
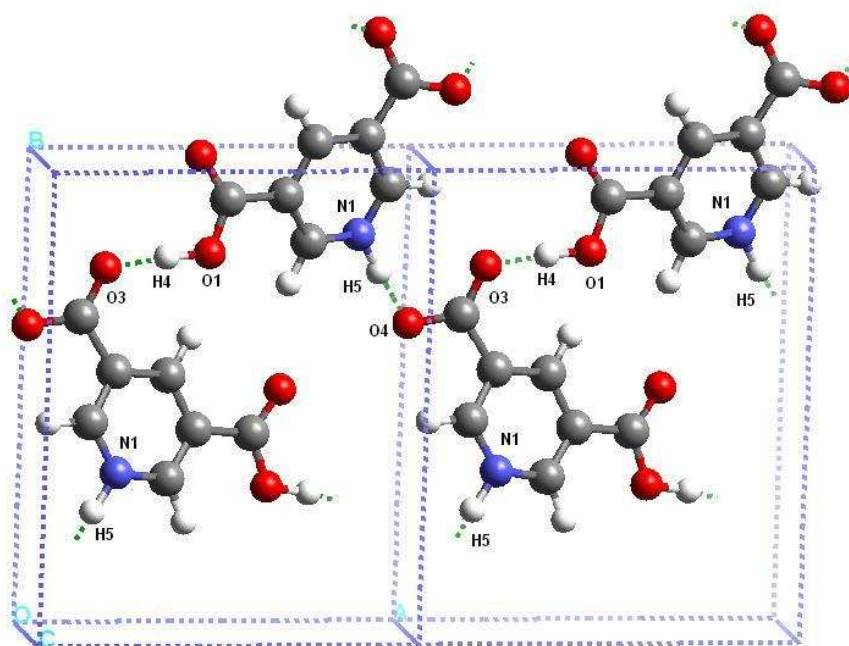
from simulation.

Figure 47

An overall view of the low temperature crystal structure of PDA as reported in [15,16].

Optimising the crystal structure with VASP leaves the *a* and *b* axes almost unchanged but causes a significant expansion in the inter-plane, *c*-direction due to the underestimate of VDW interactions in the DFT method. All simulations for PDA (and the other samples) have therefore been performed in the experimentally determined unit cells. The short hydrogen bond geometry is reasonably well reproduced with a N-O length of 2.55 Å an N-H length of 1.12 Å and an NHO angle of 176.5 degrees compared to the corresponding experimental values of 2.523 Å, 1.213 Å and 176.4 degrees respectively (see Table 5). The calculated N-H bond length is determined by the minimum in the potential energy surface whereas the experimental value takes into account the zero point energy and the wavefunction of the hydrogen atom. Analysing the energy levels in the potential well of the proton along the N-O direction (see below) gives the underestimate of the N-H bond length to be 0.03 Å. However, the "normal" O-H…O hydrogen bond becomes shorter than the N-H…O bond in terms of donor-acceptor distances, although the proton is tightly bound to the donor oxygen atom as indicated by the O-H distance of 1.05 Å (see Table 6).

C/5/1/2/ Inelastic neutron scattering and phonon calculations

The INS spectra for fully protonated PDA and a hydrogen bond deuterated sample (H4 and H5 in Figure 47), both measured on TOSCA, are shown in Figure 48 (upper curve in each panel). Also shown for each sample are the results of the corresponding VASP-PHONON calculations, only the masses of the atoms changing in the construction of the corresponding dynamical matrices.

The main effect of deuterating the two hydrogen bonds in PDA is the pronounced reduction in intensity of the band at 1400 cm^{-1}. Inspection of the eigenvectors shows the band to represent wagging motion of the N-H and O-H bonds. In the harmonic approximation, the stretching vibrations are found at the same frequency of 2200 cm^{-1} for the proton in both types of hydrogen bond in the calculation, but these are not observed in the INS spectra, even allowing for low intensity and poor statistics.

Mapping out the potential well of the proton in the N-H…O bond shows a single-well, asymmetric form, similar to the one determined for UPA (100). Numerical solutions of Schrödinger's equation for a proton in this potential give a stretch-vibration frequency of 1470 cm^{-1}. Allowing for a higher oscillator mass due to the coupling of such a low frequency, stretch vibration with other vibrational modes indicates that the main spectral change on deuteration also involves the stretching vibration.
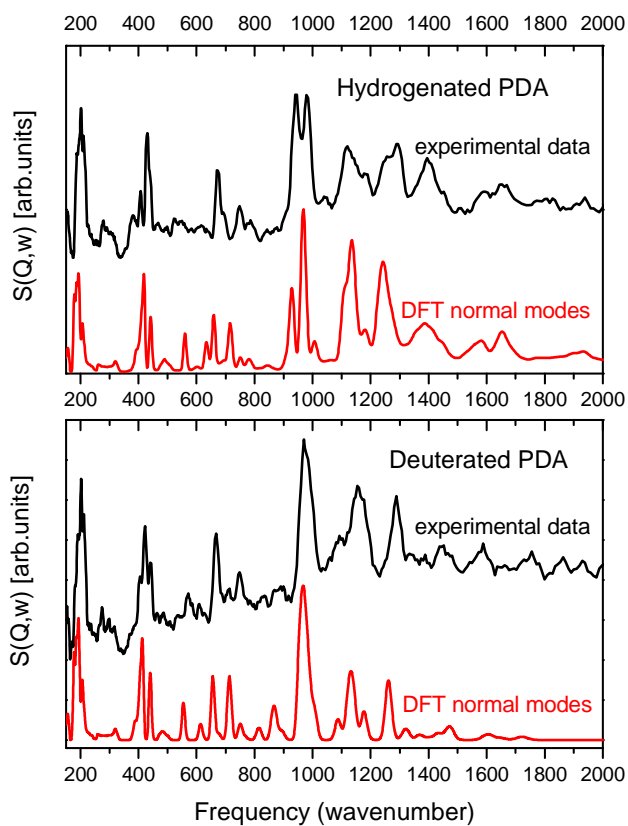
Figure 48

Comparison of measured (upper curve) and calculated (lower curve, normal modes analysis) INS S(Q,w) spectra of fully protonated (upper panel) and partially deuterated (lower panel)pyridine-3,5-dicarboxylic acid.
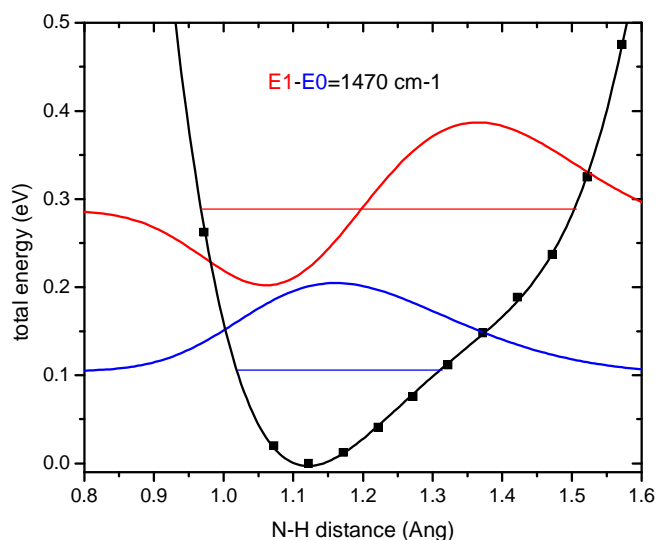

Figure 49

Potential energy well, vibrational energy levels, and wave functions of the hydrogen atom in the SSHB of the optimized low temperature structure of PDA.

C/5/1/3 Molecular Dynamics

The forces needed to determine the acceleration of atoms in a MD simulation are derived in the same way as the forces needed in a phonon calculation. Thus having validated the DFT forces against INS experimental data, the MD simulation can be expected to describe accurately the solid state dynamics of PDA. MD simulations were performed at three temperatures; 15, 250 and 300 K.

In view of the range of values discussed above for the N-H stretch frequency in the SSHB, it is useful to consider the vDOS extracted from the MD simulations for the protons in the two hydrogen bonds (Figure 50). Despite the hydrogen bonds having similar lengths in the optimised structure and at low temperature in the MD simulations (Table 5 and Table 6), their vDOS behave very differently with increasing temperature. The O-H…O bond assumes a geometry (Table 6) and vDOS of a "normal" hydrogen bond as the temperature increases with the stretching mode moving to higher frequency. On the other hand the SSHB shows the opposite behaviour. The stretching band is found at ~1800 cm^{-1} at the lowest temperature, while at 250 K and above, the band has merged with the wagging band at 1400 cm^{-1}, as indicated by the foregoing analysis of the anharmonic potential well.
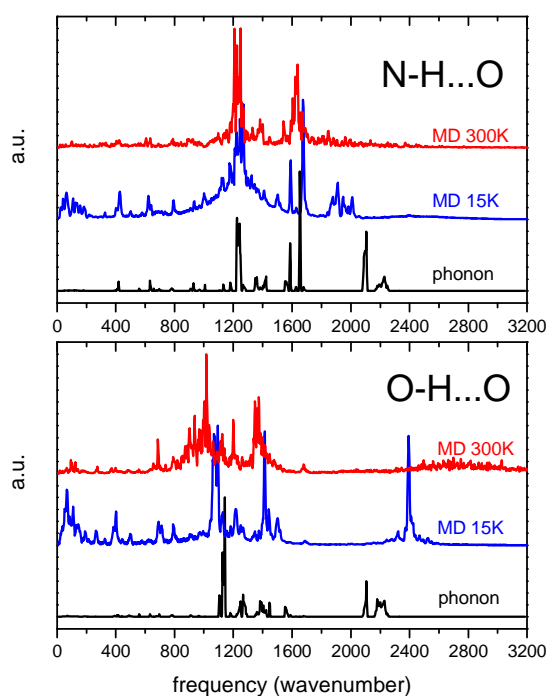
Vibrational density of states of the protons in the short N-H...O (upper panel) and "normal" O-H...O (lower panel) hydrogen bond, determined from phonon calculations (lower curves) and from MD simulations at two temperatures (upper two curves).

Extracting the donor-H distances from the MD trajectories (Table 5 and Table 6) underestimates the extent of proton transfer in the N-H...O bond, as for UPA, the average proton position moving to the centre of the hydrogen bond. However the time-dependent trace of the N-H and H...O distances shows clearly that the proton is bi-stable, spending most of its time to the left or right of centre and relatively little time at the centre. This behaviour is particularly evident from the histogram of the N-H distance over the whole MD trajectory. In contrast, the O-H...O bond, which the vDOS indicates to be a normal hydrogen bond at 250 K, shows no proton migration, the proton remaining tightly bound to the donor O atom (Figure 51, top panels).
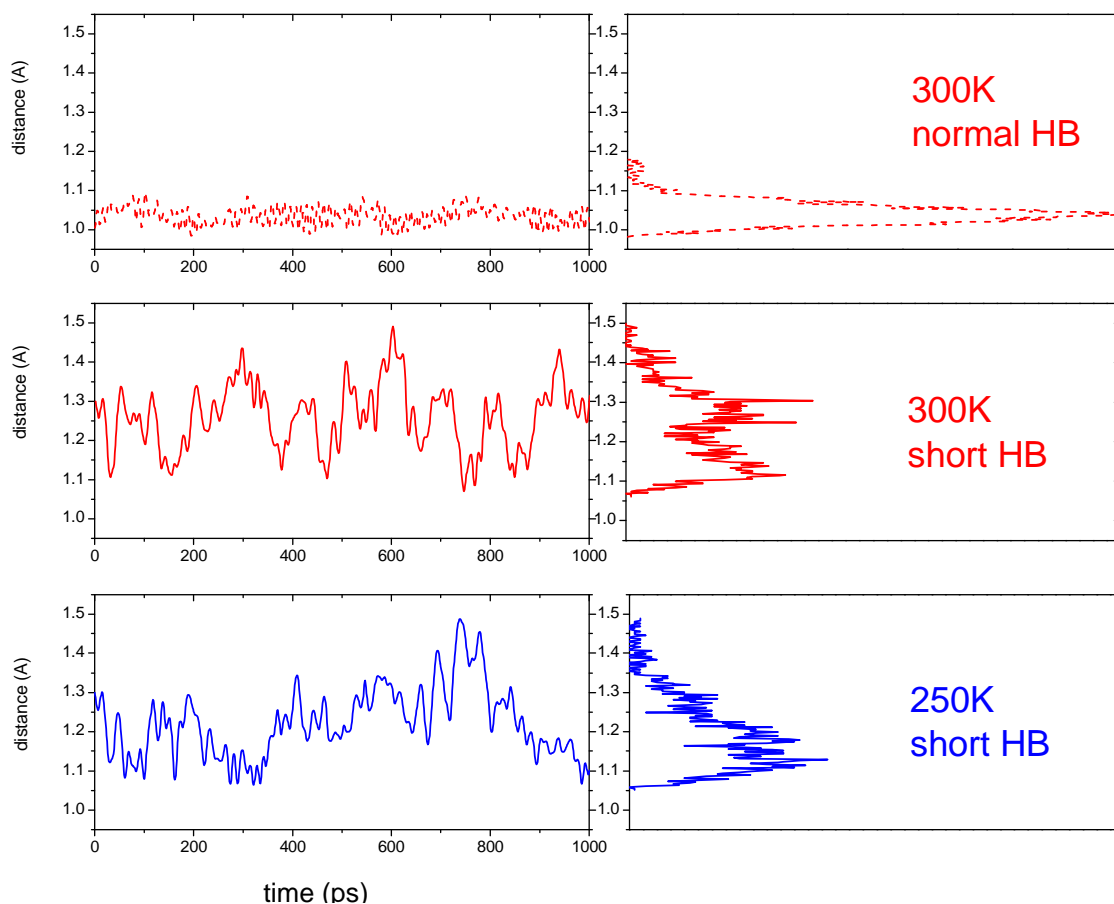

Figure 51

N-H distance (lower two panels) and O-H distance (upper panel) as a function of time over 1 ps. The histograms (right) are integrated over the whole MD trajectories.

Taking typical structures from the extrema of the 300 K trace in Figure 51 and one for the centred position of the proton in the hydrogen bond allows the corresponding potential energy wells to be calculated (Figure 52). For the extreme positions of the protons, asymmetric single wells are calculated. The centred proton position corresponds to a broad-bottomed potential well. The energy levels in these wells give rise to vibration frequencies ranging from 1200 cm^{-1} for the centred proton to 1500 cm^{-1} for the offset proton position and therefore, at ambient temperature, the N-H stretch vibration cannot drive proton migration. However the flat-bottomed potential well for the centred proton does allow the proton wavefunction to spread out and facilitate the passage of the proton from one side of the bond to the other.
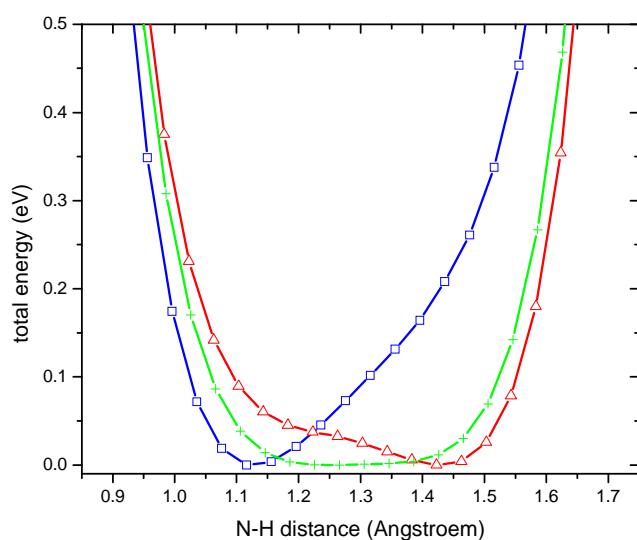

Figure 52

Potential energy well of the hydrogen atom in the SSHB of structures from MD simulation. Open square symbols; structure with hydrogen atom close to N1. Open triangles; hydrogen atom close to O4. Crosses; hydrogen atom centered.

The amplitude of the Fourier transform (FT) of the time dependent traces in Figure 51 is shown in Figure 53, which therefore shows the strength of the modulation of the N-H distance as a function of vibration frequency. The most pronounced features are a low frequency band (< 50 cm^{-1}) and peaks at 115 cm^{-1} and 145 cm^{-1}. The phonon calculations show that the first three modes are acoustic phonons, two of which are polarised in the plane. The next three modes, upto 30 cm^{-1} at the Γ -point, are out of

phase movements between the sheets, two of which are again polarised in the molecular planes. While at the Γ-point these infinitely long wavelength modes do not modulate the hydrogen bond geometry, this modulation can be achieved by shorter wavelength modes away from the Γ-point. These non Γ-point modes would account for the first low frequency peak in the FT. The next in-plane vibrations occur at 125 cm^{-1} and 155 cm^{-1}, which, allowing for an anharmonic correction, match well the peaks observed at 115 cm^{-1} and 145 cm^{-1} in the FT. Figure 54 shows the eigenvector of one of the modes at 155 cm^{-1}, which clearly modulates the geometry of the hydrogen bond.
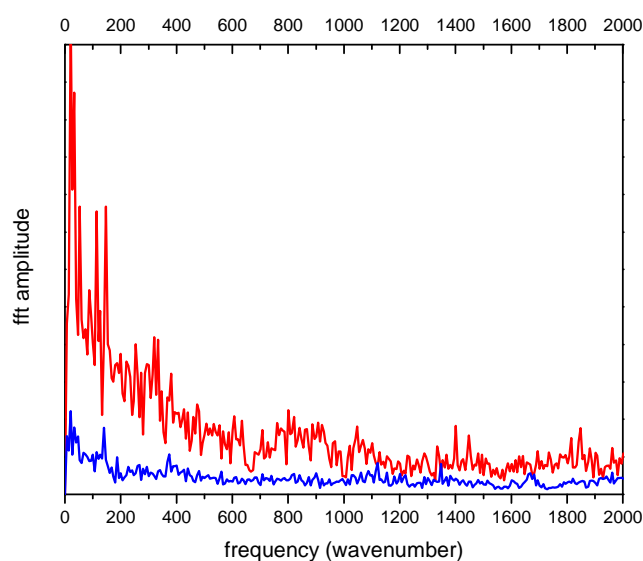

Figure 53

The Fourier Transform amplitude of the time-dependent N1-H5 bond length variation at 300K (upper curve), and 15K (lower curve).
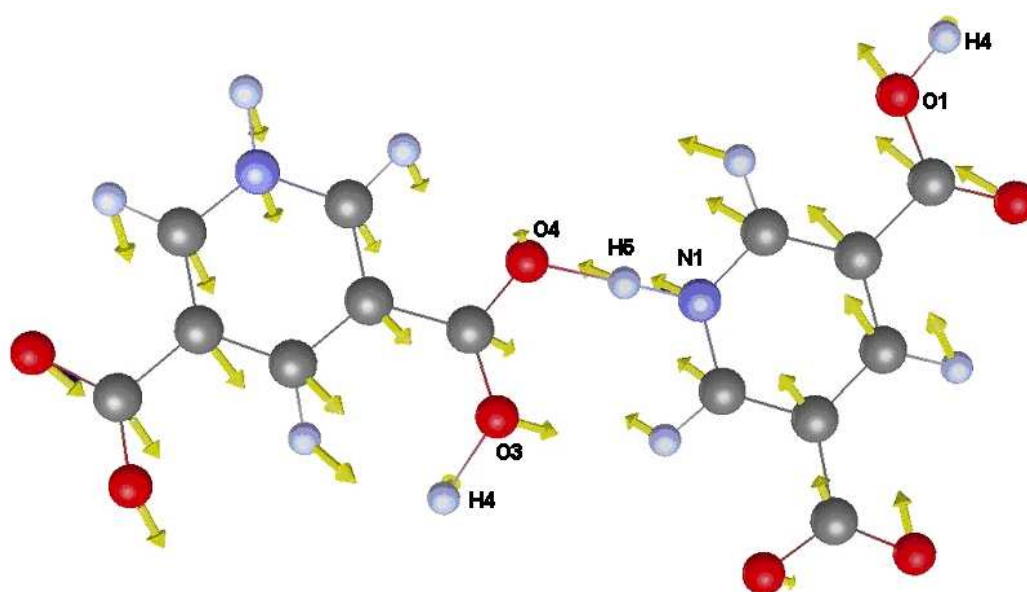

Figure 54

Representation of the mode at 155 cm^{-1}.

Partially deuterated PDA presents, experimentally, a more pronounced proton transfer effect, the smallest zero point motion tending to localize the proton nearer the donor or acceptor atom. As mentioned above, zero point motion is not included in our MD simulations and calculated proton transfer in the partially deuterated compound is very similar to that of the fully protonated compound (compare Table 5 and Table 7). Figure 51-Figure 54 are almost unchanged for partially deuterated PDA. While low frequency molecular and lattice vibrations are almost unchanged upon deuteration, this is not the case of the higher frequency modes involving the hydrogen bond protons. The foregoing vibrational analysis indicates that stretching and wagging modes of the N-H bond occur at 1400 cm^{-1} for N-H and therefore at ~990 cm^{-1} for N-D, contributing weakly to the strongest band in the INS spectra. If high frequency vibrations played a role in proton transfer then the strong variation of these frequencies upon deuteration would lead to markedly different proton dynamics. That similar proton transfer effects are observed for protonated and deuterated compounds is further evidence of the dominant role of low frequency lattice dynamics.

	N1....O4	N1....D5	D5....O4	N1-D5-O4
exp data				
296K	2.564	1.457	1.108	177.6
15K	2.538	1.151	1.389	175.9
MD simulations				
300K	2.553	1.251	1.312	169.9

Table 7

Proton transfer as a function of temperature in the N1...D5...O4 SSHB in deuterated PDA as reported in [16].

C/5/2/ 1:2 co-crystal of benzene-1,2,4,5-tetracarboxylic acid and 4,4'-bipyridyl

C/5/2/1/ Structural information

The structure of the 1:2 co-crystal of benzene-1,2,4,5-tetracarboxylic acid and 4,4'-bipyridyl was reported at a number of temperatures in the range 20K-296K by Cowan et al [98]. BTA-BPY crystallizes in a triclinic structure, in the P-1 space group with two molecules in the unit cell. The SSHB N1-H1…O22 is shown in Figure 55 with atom labeling consistent with that of Cowan et al. The BTA and BPY molecules form hydrogen-bonded, 1D chains.

	N1….O22	N1….H1	H1….O22	N1-H1-O22
exp data				
296K	2.531	1.302	1.241	169.6
200K	2.521	1.251	1.281	169.4
20K	2.522	1.207	1.325	170.2
MD simulations				
300K	2.519	1.224	1.313	164.5
200K	2.534	1.141	1.401	164.9
20K	2.541	1.106	1.446	169.8

Table 8

Proton transfer as a function of temperature in the N1…H1…O22 SSHB in BTA-BPY as reported in [98]; experimental data and simulation results.
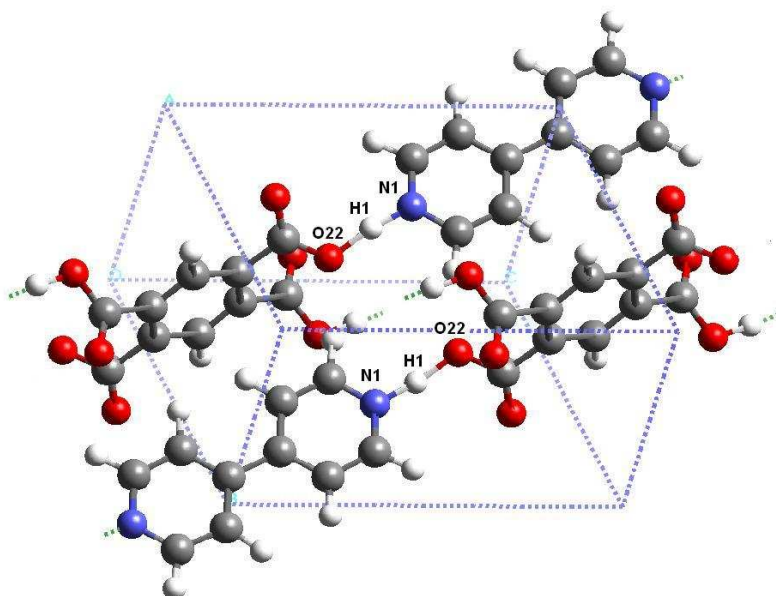

Figure 55

Low temperature crystal structure of the 1:2 co-crystal of benzene-1,2,4,5 tetracarboxylic acid and 4,4'-bipyridyl as reported in [98]

As with UPA and PDA the hydrogen bond length is overestimated by simple geometry optimisation and the low temperature MD run (Table 8). Another characteristic of the DFT calculations is the donor-H distance which is shorter by 0.1 Å than in the measurement. The evolution of the hydrogen bond geometry with temperature is very similar to that of PDA (compare Table 5 and Table 8). The time-dependence of the N-H and H..O distances (Figure 56) also hints at the same bi-stable behaviour, which can be better detected by projecting the traces onto a histogram (Figure 56).

The Fourier transform of the time-dependent traces in Figure 56 shows, as for PDA, that the low frequency phonons give rise to the strongest modulation of the N-H bond length (Figure 57), although, in this case, there are no well-defined frequencies.

C/5/3/ 1:1 crystal of 4-methylpyridine and pentachlorophenol

C/5/3/1/ Structural information

The structure of the 1:1 crystal of 4-methylpyridine and pentachlorophenol was reported at low temperature by Steiner et al.[99]. 4-MP-PCP crystallizes in triclinic structure, in the P-1 space group with two molecules in the unit cell. The SSHB O-H…N1 is shown in Figure 58 with atom labeling consistent with that of Steiner et al. The two molecules in this crystal form hydrogen-bonded dimers.

	N1….O	N1….H	H….O	N1-H-O
exp data				
200K	2.525	1.304	1.229	170.6
100K	2.513	1.265	1.258	170.1
20K	2.506	1.206	1.309	170.4
MD simulations				
300K	2.562	1.136	1.441	166.7
200K	2.561	1.112	1.462	167.8
20K	2.616	1.11	1.521	168.7

Table 9

Proton transfer as a function of temperature in the N1…H…O SSHB in 4-MePy-PCP as reported in [99]; experimental data and simulation results.

Table 9 shows that at low temperature the calculated hydrogen bond length is overestimated to a much greater extent (0.1 Å) than in the other samples. As a result there is little evolution in the hydrogen bond geometry as a function of temperature, the donor-acceptor distance shortening by 0.05 Å which is mainly accomodated by a reduction in the proton-acceptor distance.
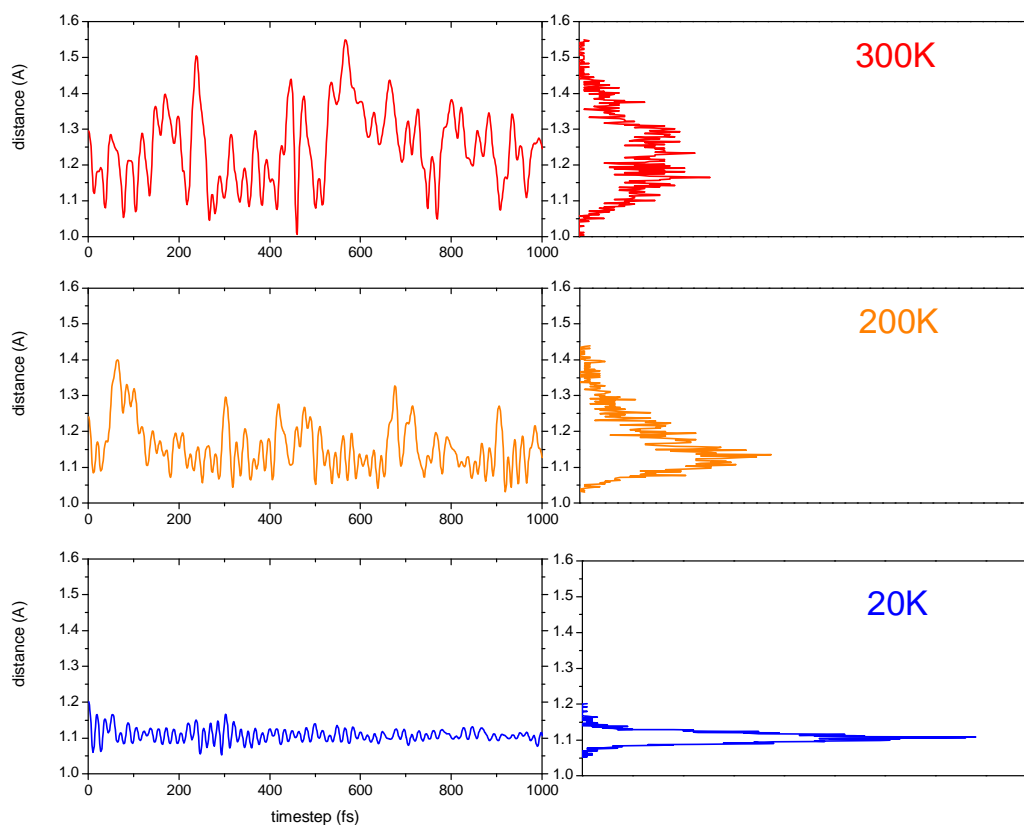

Figure 56

Geometries obtained for the short hydrogen bond N1...H1...O22 in BTA-BPY by MD simulation at three temperatures. Histograms integrated over the full trajectories are shown on the right.
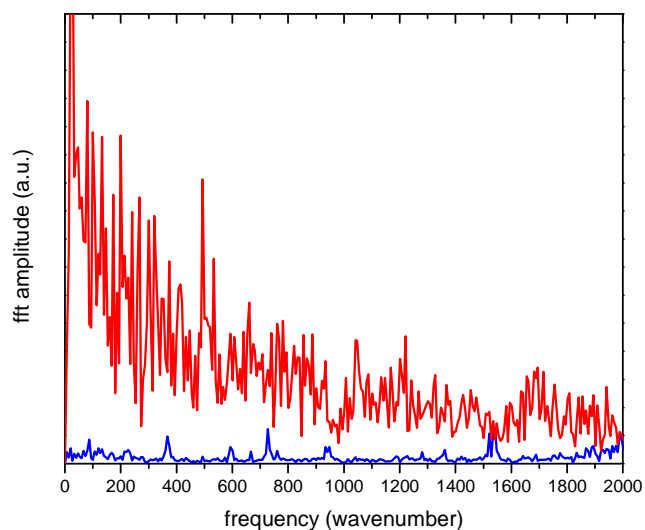
Figure 57

The Fourier Transform amplitude of the time-dependent N1-H1 bond length variation at 300K (upper curve), and 15K (lower curve).
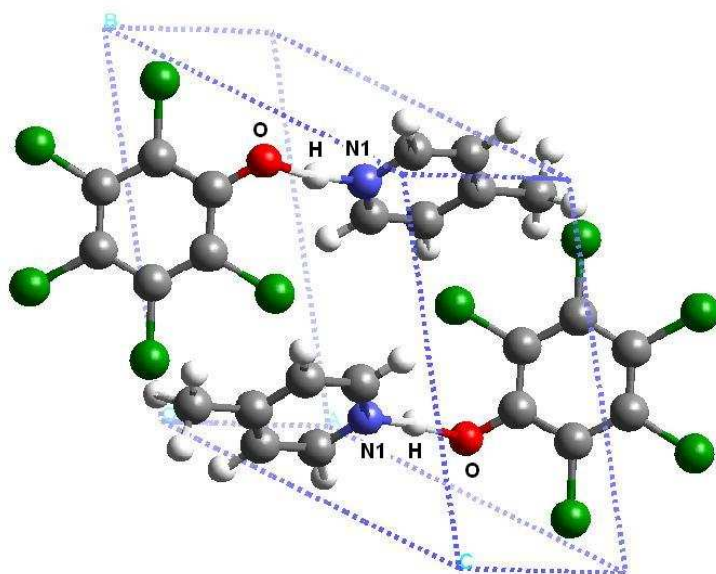

Figure 58

Low temperature crystal structure of the 1:1 crystal of 2-methylpyridine and pentachlorophenol as reported in [99]

C/6/ Discussion

Using INS and ab initio MD, we have investigated three more hydrogen-bonded systems that show proton migration as a function of temperature. In these systems the

SSHB are of the type N-H…O, compared to UPA which has a O-H…O SSHB. Proton migration is therefore not chemically limited to a particular type of hydrogen bond.

Neutron diffraction provides the reference crystallographic data, which is the basis of all simulations presented here. As in the case of UPA, solid state DFT methods tend, in geometry optimisation and low temperature MD calculations, to overestimate the donor-acceptor distance and underestimate the donor-H distance, which hinders the proton migration effect in the simulations. This trend is most pronounced in the case of 4MP-PCP to the extent that there is no evidence of proton migration in the MD simulations. In 4MP-PCP the hydrogen bonds connect the two molecules in dimers, whereas all other systems studied present infinite chains (BTA-BPY), 2D sheets (PDA) or 3D networks (UPA). Infinite networks of hydrogen bonds provide long range constraints that do not exist for a dimer and in this case the SSHB have more freedom to lengthen on geometry optimisation.

In all other cases, the extent of proton migration in MD simulations is less than in the corresponding measurements. Increasing the temperature above that of the experiments helps to an extent but does not allow the full amplitude of migration to be reproduced. In addition to the problems mentioned above of SSHB geometry, the fact that the protons are treated as classical, point particles instead of partially delocalised wavefunctions also hinders proton migration. Calculation of the eigenfunctions in the potential well of the proton along the donor-acceptor direction indicate that neglecting the zero point motion amounts to a shift ~0.03 Å of the proton away from the centre of the SSHB, compared to a transfer distance of ~0.1 Å.

One significant difference between the N-H…O SSHBs studied here and the O-H…O SSHB in UPA is that, experimentally, the protons in all three cases migrate beyond the centre of the hydrogen bond. In BTA-BPY evidence of this behaviour is seen in the histogram of proton positions (Figure 56) and the time dependent traces of the N-H distance which clearly show the proton spending a significant amount of time beyond the centre of the hydrogen bond. This effect is much clearer in PDA. The time dependence of the N-H distance shows almost periodic hopping of the proton from one side of the hydrogen bond to the other and the histogram clearly shows two peaks separated by ~0.1 Å as in the experiment.

Fourier transforming the time domain information shows that, as in the case of UPA, low frequency lattice vibrations are driving proton migration. For BTA-BPY, the time dependence of the N-H distance does not show any obvious modulation and there are no clear peaks in the FT. For PDA, there is almost periodic hopping across the centre of the hydrogen bond (period 100 – 200 fs) and the FT (Figure 53) shows an initial peak at low frequency (< 50 cm^{-1}) and two well-defined peaks at 115 and 145 cm^{-1}. INS and phonon calculations show that these features correspond to vibrations in the plane of the molecular sheets which give rise to the strongest modulation of the hydrogen bond geometry.

Analysing structures for PDA from the MD trajectories show that the lattice vibrations tend to stabilise the proton position on either side of the SSHB rather than in the centre. The corresponding potential wells are single-well and asymmetric for the offset proton positions and flat-bottomed for the central position. In a model of a single particle oscillator of one proton mass, the calculated vibrational frequencies are lowest for the flat-bottom potential with a limiting value of 1200 cm^{-1}. These N-H stretching movements, coupled in reality to other high frequency vibrations, can never play a role in proton migration which occurs at temperatures lower than 300 K.

The apparent bi-stable behaviour of SSHBs in which protons migrate beyond the centre of the hydrogen bonds has not been detected experimentally. The time dependent traces and histograms shown here are in the reference frame in which the donor N atom is stationary. Calculating the same quantities in the reference frame centred at the origin of the crystallographic cell gives rise to broad, featureless histograms. Diffraction measurements only have access to these broad distributions. Spectroscopic measurements would therefore be required to probe a two-site jump process as have been performed for quantum tunnelling between the two wells of the potential of a normal hydrogen bond [102]. Whereas these quasielastic measurements were performed with μeV resolution, that is on the nanosecond timescale, only meV resolution would be required to detect the fast hopping process predicted here.

C/7/ Conclusion

We have extended initial work on proton migration in SSHBs in UPA to three more systems which show proton transfer. If DFT methods allow hydrogen bond geometries to be accurately determined (two out of three cases studied here) then MD simulations show the same effect as seen in neutron diffraction experiments, albeit to a lesser extent. The two systems for which the simulations are successful reveal the same mechanism for proton migration as in UPA, that is lattice vibrations stabilise solid state structures at non-zero temperatures for which the proton is partially or completely transferred. When a proton transfers completely from donor to acceptor as in PDA and BTA-BPY, a bi-stable regime exists at intermediate temperature where the proton avoids a relatively unstable, centred position. In the case of PDA, specific phonons which modulate the SSHB geometry are seen to be responsible for proton transfer.

While many systems have SSHBs, only a few show proton migration behaviour. Understanding this type of behaviour requires a precise knowledge of the low frequency dynamics of the system, which has been achieved here for PDA with INS experiments and phonon calculations. Having identified specific phonons that drive proton migration it would be useful to characterise the lattice dynamics of SSHB systems that don't show proton migration. In view of the key role of dynamics in proton migration predicting the activity of SSHBs in complex systems is clearly very challenging.

Appendix D

Neutron vibrational spectroscopy gives new insights into the structure of poly (p-phenylene terepthalamide)

M. Plazanet, F. Fontaine-Vive, K.H. Gardner, V.T. Forsyth, A. Ivanov, A.J. Ramirez-Cuesta and M.R. Johnson, J. Am. Chem. Soc. 127, 6672 (2005)

D/1/ Abstract

The vibrational spectra of benzanilide and poly(p-phenylene terephthalamide) have been measured using inelastic neutron scattering. These compounds have similar hydrogen bond networks, which, for poly(p-phenylene terephthalamide), lead to two-dimensional hydrogen-bonded sheets in the crystal. Experimental spectra are compared with solid state, quantum chemical calculations based on density functional theory (DFT). Such "parameter-free" calculations allow the structure-dynamics relation in this type of compound to be quantified, which is demonstrated here for benzanilide. In the case of poly(p-phenylene terephthalamide), vibrational spectroscopy and DFT calculations help resolve long-standing questions about the packing of hydrogen bonded sheets in the solid state.

D/2/ Introduction

In materials lacking long-range structural order vibrational spectroscopy is a powerful method for obtaining structural information. Diffraction techniques rely on long-range order to give well-defined Bragg reflections, whereas the vibrational modes of molecules depend on the local intra and inter-molecular environment. Inter-molecular interactions typically extend over distances of tens of Å, but the effective range of interactions that determine vibrational modes in organic compounds is more like 5-6 Å. The structure-dynamics relation can be established empirically as is the case for determining the presence of secondary structures like β-sheets and helices in proteins from the optical, spectral profiles of the amide bands [103,104]. Indeed the study of β-sheets has assumed greater improtance over the last decade since highly aggregated β-sheets have been established as symptoms of amyloid and prion diseases. However, a higher level of structural information is potentially available if a quantitative link can be made between structure and vibrations via accurate inter-atomic potentials. In the

case of polypeptides, first principles-based calculations of spectral profiles have been employed, although these have been limited to the amide bands [105,106].

One of the original uses of vibrational spectroscopy was precisely that of parametrising inter-atomic interactions. Typically spectral frequencies rather than intensities were used for this purpose since, for optical spectroscopies (IR and Raman), spectral intensities have been difficult to determine, a knowledge of dipole moments and polarisability, respectively, being required. Although these quantities for optical spectroscopy can now be extracted from quantum chemical calculations, in inelastic neutron scattering (INS) spectral intensities are directly related to vibrational amplitudes via tabulated scattering cross-sections. The information content of INS spectra is more easily and reliably exploited.

The quantitative link between structure and dynamics depends on the accuracy of the inter-atomic potentials that are used to either calculate the vibrational density of states from molecular dynamics simulations, or the normal modes from the dynamical matrix. For smaller systems (less than a few hundred atoms) the second approach entails fewer calculations and gives directly the vibrational modes, including their wavevector dependence (phonon dispersion). This approach has been exploited in recent years to study hydrogen-bonded molecular crystals, using "parameter-free", solid state, density functional theory (DFT) based methods to determine the inter-atomic force constants [3,4,5,6,7].

The purpose of this work is therefore two-fold: to extend recent experimental and computational work to hydrogen bonded compounds more closely related to secondary structures in biological molecules and to investigate the solid state structure-vibration relation in a crystalline polymer. Poly(p-phenylene terephthalamide) (PPTA) is a polymer which crystallises in the form of hydrogen-bonded sheets, amide linkages in the polymer allowing hydrogen bonding between polymer chains to form 2D sheets. PPTA is a highly crystalline polymer that can be spun into fibres which exhibit exceptional thermal and mechanical properties [107]. It has numerous commercial applications and is sold under the trade names of Kevlar® and Twaron®. The fibre used in this work is estimated to have a have a crystallinity better than 90%, but the crystallites are small, ~50 Å in diameter perpendicular to the fibre axis [108], limiting the long-range order and resulting in about 30% of the

polymer chains being on the surface of crystallites. Thus vibrational spectroscopy is an important tool for structure investigation. Uncertainty persists concerning the packing of hydrogen-bonded sheets in the crystal structure of PPTA, which is governed by weak intermolecular interactions. While the small crystallite size and partial crystallinity of PPTA degrade the quality of vibrational information, the fibrous nature of the sample allows additional information to be obtained by orienting it in the neutron beam and measuring the polarisation of the vibrational modes directly.

Optical spectroscopy techniques have been widely applied to PPTA in the past. IR spectra have been used to identify solid state structures resulting from two different production techniques and the vibrational modes were tentatively assigned based on atom-atom potentials [109]. Force field parameters were further refined (as described above) based on the structure of the hydrogen-bonded sheet and then used to investigate mechanical properties of fibres [110]. Total energy calculations, based on empirical force fields, have been used to investigate the relative stability of two structures of PPTA [111], but the structure that will be shown here to be the most plausible was not studied. It will be demonstrated below that the energy differences between different solid state structures are small and the precision of first principles calculations is required. The same precision is needed for accurate determination of vibrational spectra.

In order to validate the DFT based methods used to establish the structure-dynamics relation for PPTA, we, like others [112], have also studied, as a model system, a highly crystalline powder of benzanilide, which is approximately the monomer of PPTA and also forms hydrogen-bonded chains in the crystal.

In the following section we present in detail the structural information relevant to the samples in this work. In section III, computed and measured spectra are compared for the two compounds. The connection between structure and vibrations and the limitations of the approach adopted here are discussed in section IV. The conclusion and perspectives for vibrational spectroscopy in biomaterials is presented in section V.

One of the original uses of vibrational spectroscopy was precisely that of parametrising inter-atomic interactions. Typically spectral frequencies rather than intensities were used for this purpose since, for optical spectroscopies (IR and Raman), spectral intensities have been difficult to determine, a knowledge of dipole moments and polarisability, respectively, being required. Although these quantities for optical spectroscopy can now be extracted from quantum chemical calculations, in inelastic neutron scattering (INS) spectral intensities are directly related to vibrational amplitudes via tabulated scattering cross-sections. The information content of INS spectra is more easily and reliably exploited.

D/3/ Structural Information

The structure of PPTA was first determined by Northolt [40,41], an orthorhombic structure in the monoclinic space group *Pc* (*P11n*) being proposed. Details of the lattice are given in Table 10. Each polymer chain in the unit cell forms part of a 2D structure in which neighbouring chains are connected at the amide linkages by hydrogen bonds (Figure 59). The hydrogen-bonded sheets are separated by ~4 Å (*a/2*). The relative alignment between sheets is characterised by the phenyl groups and the amide linkages. In the Northolt structure, terephthaloyl residues are offset by ~6.5 Å (~*c/2*), all amide linkages are oriented approximately along the *b*-axis and these groups have a similar, parallel alignment in neighbouring sheets. A similar structure but with higher symmetry (*P2₁/n*), which enforces the *c/2* offset of equivalent phenyl groups, was proposed by Tadokoro and co-workers [43]. Later, Liu and co-workers concluded from electron diffraction data from single crystals [44] that the correct space group is *Pa*. In this structure, terephthaloyl residues are not offset and the amide linkages have an anti-parallel configuration along the *b*-axis between neighbouring sheets. A structure similar to that of Liu can be obtained from the Northolt structure by introducing a ~6.5 Å relative translation of neighbouring sheets along the polymer axis, the essential difference being that the Liu structure presents a herring-bone packing of the phenyl rings, whereas the "Northolt-translated" structure has a parallel packing of the phenyl groups. Finally, a recent, neutron fibre diffraction experiment, in which the terephthaloyl residues were selectively deuterated, has confirmed the zero offset along *c* between equivalent residues in neighbouring sheets [42]. From the

initial analysis of the neutron diffraction data, a second structure in the same space group (*Pb*), in which the amide linkages have a parallel alignment between neighbouring sheets, could not be ruled out. Figure 60 shows the four structures discussed above (Northolt, Northolt-translated, Liu (*Pa*) and *Pb*) and the relevant structural information is presented in Table 10. It should be noted that these four structures have a common hydrogen-bonded sheet structure, the packing of these sheets giving rise to the different, solid state structures.

Benzanilide crystallises in the monoclinic space group Cc (Z=4) [113]. This structure (Figure 61) closely mimics that of Liu for PPTA in terms of hydrogen bonding, the relative orientations of amide linkages and the relative positions of phenyl groups, the main difference being the absence of polymeric chains due to the absence of a second amide linkage on the benzanilide molecule.

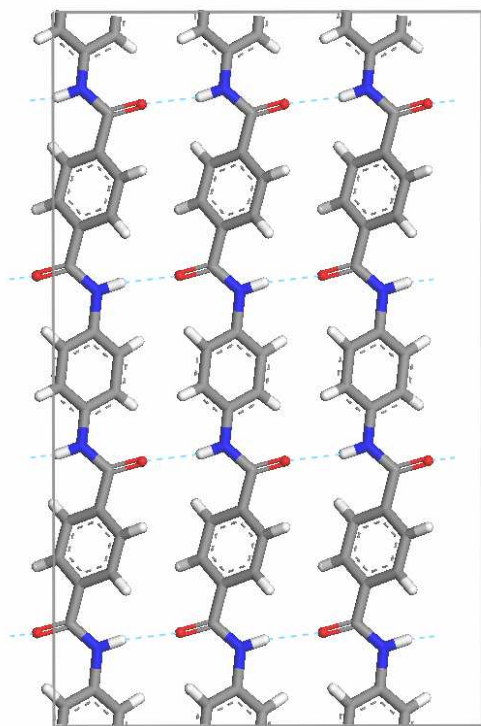|	Northolt	Northolt translated	Liu	*Pb*
a [Å]	7.87	7.87	7.88	7.88
b [Å]	5.18	5.18	5.22	5.22
c [Å]	12.9	12.9	12.9	12.9
α, β, γ	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Space group	P 11n	P 11n	P 1a1	Pb11
Chain location [x,y]	[0,0],[1/2,1/2]	[0,0],[1/2,1/2]	[1/4,1/4],[3/4,3/4]	[1/4,1/4],[3/4,3/4]
Z	2	2	2	2
Glide plane symmetry operator	x+1/2, y+1/2, -z	x+1/2, y+1/2, -z	x+1/2, -y, z	-x, y+1/2, z
c translation between sheets	~0.5	~0	0	0

Table 10

Structural information for PPTA

Figure 59

Two dimensional hydrogen-bonded sheet structure common to all PPTA structures. The N-H...O=C hydrogen bonds are shown by dashed lines.

Northolt (*Pn*)   Northolt-translated (*Pn*)   Liu (*Pa*)   *Pb*
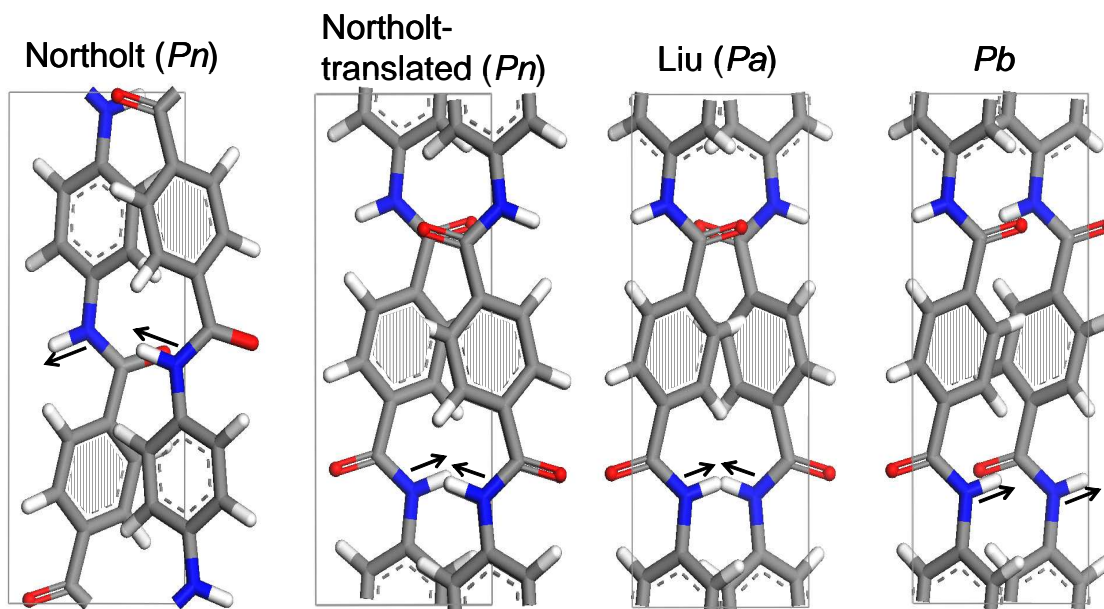
Figure 60

The principle PPTA structures studied in this paper. The two chains in the unit cell are shown, that is one from each of the neighbouring sheets. Terephthaloyl residues are shaded to show the alignment of the chains. Arrows on the N-H bonds indicate the orientations of the amide linkages. The Northolt and Pb structures have an approximately parallel alignment of amide linkages, the Northolt-translated and Liu structures have an anti-parallel alignment. The packing of phenyl rings between sheets in all structures is of herring-bone type, except in the Northolt-translated structure in which the phenyl groups are stacked with a similar, parallel orientation.
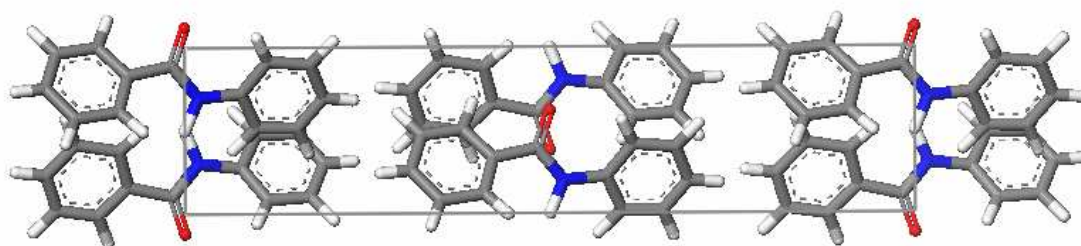
Figure 61

The crystal structure of benzanilide (ab plane), space group Cc, cell; a = 24.34 Å, b = 5.33 Å, c = 8.01 Å, β = 107.2°

D/4/ Vibrational Spectroscopy

D/4/1/ Benzanilide

Figure 62 shows the experimental INS spectrum for the powder sample of benzanilide (measured on the TOSCA spectrometer at the ISIS facility, UK [75]) and the calculated, Gamma point, one-phonon spectrum. The spectral intensity is almost entirely due to incoherent scattering from hydrogen. The broad background intensity in the experimental spectrum is due mainly to multi-phonon excitations. In view of the fact that the calculated spectrum contains 312 modes (three times the number of atoms in the unit cell), the agreement between these spectra is very good and serves to validate the use of solid state DFT methods (as implemented in the VASP code [88,89,90,91]) to determine molecular vibrations in this type of hydrogen-bonded solid, with only the crystal structure as input. The most apparent discrepancies between measurement and calculation relate to small errors in mode frequencies, which result in structured, calculated bands instead of intense peaks due to degenerate

modes (e.g. at ~400 cm^{-1}) and vice-a-versa (e.g. at ~900 cm^{-1}). With respect to the case of PPTA (to be discussed below) we draw attention to the out-of-plane N-H wagging vibrations, which are coupled to similar, out-of-plane vibrations of C-H bonds on the neighbouring phenyl rings, and are found in a family of nine modes between 667 cm^{-1} and 710 cm^{-1}. The relatively high frequency of these modes is due to the hydrogen bond network. The intense, calculated peak just below 700 cm^{-1} would therefore be expected to be broadened by dispersion due to the hydrogen bond network.

In view of the accuracy of the Gamma point calculation, the vibrational density of states was not calculated as this would have required computationally intensive calculations on a supercell composed of two unit cells (a,2b,c, volume ~1984 Å^{3}), the VASP calculations scaling approximately with the cube of the number of atoms.
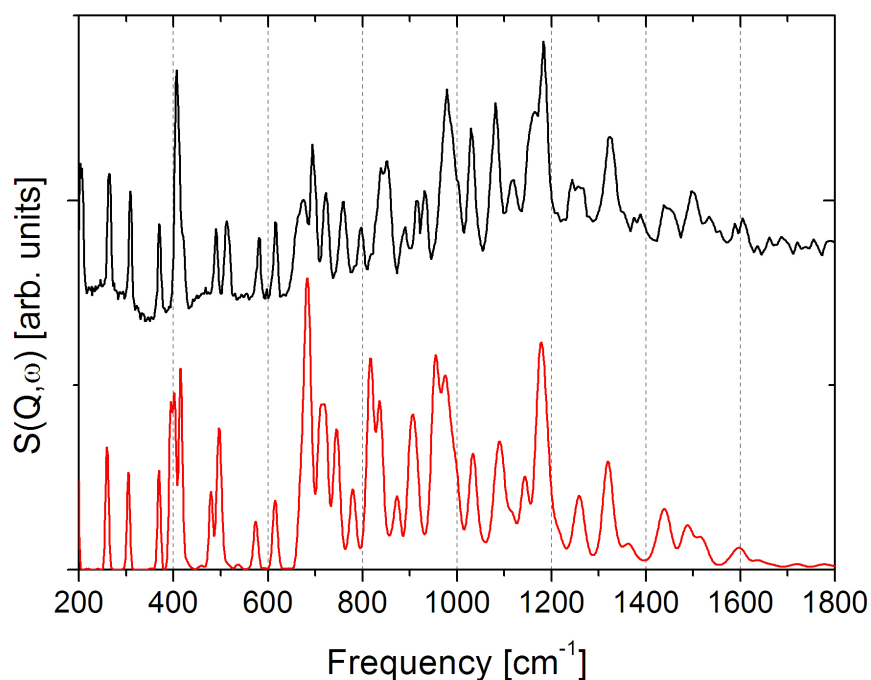
Figure 62

The INS spectrum of benzanilide measured on TOSCA (top) and calculated for the gamma point using VASP and CLIMAX (bottom).

D/4/2/ PPTA

Ten INS spectra of PPTA were measured on the IN1 spectrometer (at the ILL, France [76]) at a temperature of 10 K, as a function of fibre orientation with respect to the incident beam (approximately the momentum transfer vector **Q**) ranging from parallel to perpendicular. These spectra were corrected for the effective, illuminated area of the sample and summed to give the upper spectrum in Figure 63 (curve e). Modes above 1400 cm^{-1} could not be measured due to the lack of hot neutrons on IN1 at the time of the experiment. The rising background is due, in part, to multi-phonon excitations and an orientation-dependent background signal. A couple of these measurements were repeated on the higher resolution, TOSCA instrument, but this did not result in significantly more detailed spectra, effective resolution being limited by the crystallite size and partial crystallinity of the sample.

In view of the smaller size of the unit cell, the vibrational density of states, rather than the Gamma point modes, was calculated for the four PPTA structures shown in Table 10 and Figure 60 from supercells of dimension 8 Å x 10 Å x 12 Å, composed of two unit cells. The bottom panel (a) in Figure 63 shows the calculated, summed spectrum for the original Northolt structure. The total scattering function has been convoluted with the frequency dependent, instrument resolution function to give the smooth curve. The lower curve in the panel is the partial scattering function (not convoluted) due to the hydrogen atom in the amide linkage, showing the contribution of this atom to the spectrum. Despite the lower resolution of the PPTA spectra, the agreement between calculation and experiment is much worse than in the case of benzanilide. In particular, the triplet of peaks (690, 830, 950 cm^{-1}) in the intermediate frequency range in the experimental spectrum is not well-reproduced. The strong peak at 790 cm^{-1} in the calculated spectrum has a dominant contribution from pure, N-H out-of-plane, wagging modes.

In the spectral range of Figure 63, the amide linkage hydrogen atom also contributes intensity at ~1250 cm^{-1} through N-H in-plane wagging modes. Beyond this spectral window, the calculations predict low intensity bands at ~1600 cm^{-1} due to C=O stretching modes and at 3200 and 3400 cm^{-1} due to C-H and N-H stretching modes respectively. Such a high N-H stretching frequency indicates that the hydrogen bond is weak, the optimised Northolt structure giving a N-H bond length of 1.04 Å, a N…O distance of 2.98 Å and a N-H…O angle of 165°.

Since the principal result from the recent neutron diffraction experiment was that equivalent phenyl rings should not be offset along the direction of the polymer chains between neighbouring sheets, the "Northolt-translated" structure (described above) was investigated. The calculated spectrum (Figure 63, panel b) is in better agreement with the measured spectrum, the amide group hydrogen atom contribution being split between the peaks at 740 cm^{-1} and 810 cm^{-1}.

Calculations were performed on the *Pb* structure that was considered to be compatible with the initial inspection of the fibre diffraction data. The spectrum (Figure 63, panel c) strongly ressembles that of the Northolt structure (panel a), again pure N-H wagging modes making a significant contribution to the most intense peak (at 755 cm^{-1}). The common feature of the *Pb* and Northolt (*Pn*) structures is the parallel configuration of amide linkages between sheets.

Finally, calculations on the Liu (*Pa*) structure, which is compatible with the fibre diffraction data, give the spectrum shown in Figure 63, panel d. This spectrum is in best agreement with the experimental data. Virtually all calculated peaks are close to the measured peak positions and variations from peak to peak, in intensity and profile, are well-reproduced. In contrast to the other structures, especially Northolt and *Pb*, the N-H wagging modes make a relatively weak contribution to the spectrum since they are strongly mixed with the out-of-plane C-H wagging modes. Like the Northolt-translated structure, the amide linkages have an anti-parallel alignment. The difference between the two structures concerns the packing of the phenyl groups, which participate in many vibrations and therefore lead to small differences throughout the spectrum.

The spectra presented in Figure 63 have all been summed over a range of fibre orientations with respect to **Q**. Information about the polarisation of the vibrational modes, and further validation of the calculation, can be sought from the orientation dependence of the spectral intensities, since modes give the strongest intensity when **Q** is parallel to the principal direction of atomic displacements. Figure 64 shows the measured spectra and those calculated for the Liu structure, the fibre axis is parallel to **Q** in the top spectrum and perpendicular in the bottom spectrum. The measured and calculated variations in intensity with fibre orientation match well. Most modes in the range from 200 to 1000 cm^{-1} are polarised perpendicular to the fibre axis, like the out-

of-plane N-H wagging modes. Only the band at 1100 to 1300 cm^{-1} has a significant longitudinal polarisation, these higher frequency modes involving wagging of the C-H bonds in the planes of the phenyl groups and wagging of the N-H bonds in the planes of the amide linkages. The decay in intensity of these modes is smoother in the experimental spectra than in the calculated spectra because the detector on IN1 covers a 15° 2θ scattering angle around 90°, whereas the calculation is performed for a single scattering angle of 90°.
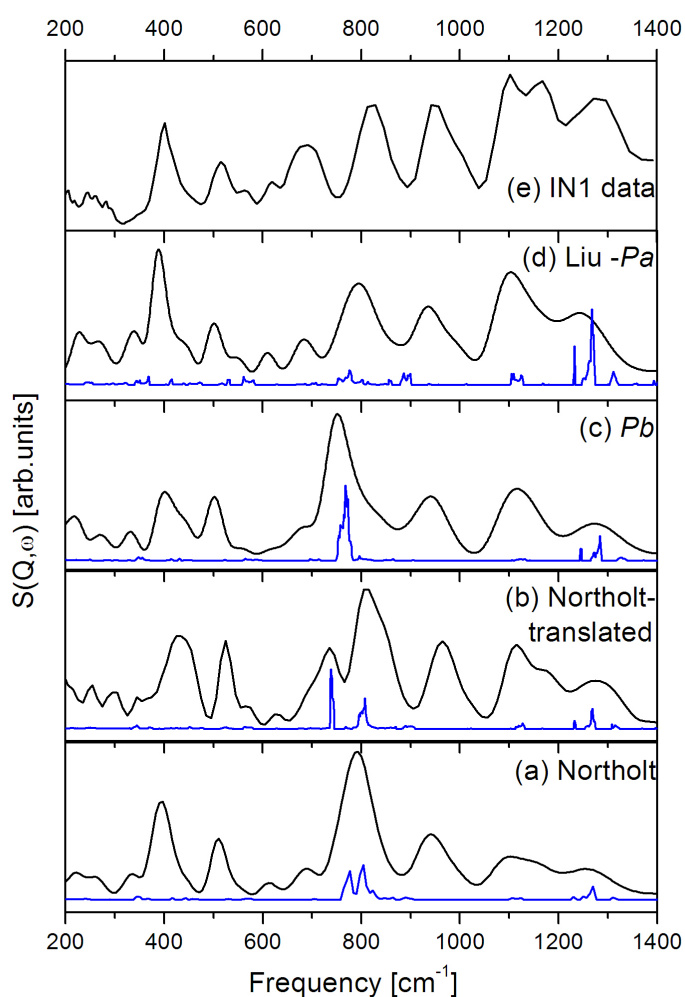
Figure 63

The measured INS spectrum of PPTA, summed over fibre orientations ranging from parallel to the momentum transfer Q vector to perpendicular (top, e). Calculated INS spectra for the same, average fibre and instrument configuration are shown in the lower four panels. The smooth curves are the result of convolution of the calculated total scattering function with the frequency dependent, instrument resolution function.

The unconvoluted, partial scattering function for the amide hydrogen is also shown in each panel.
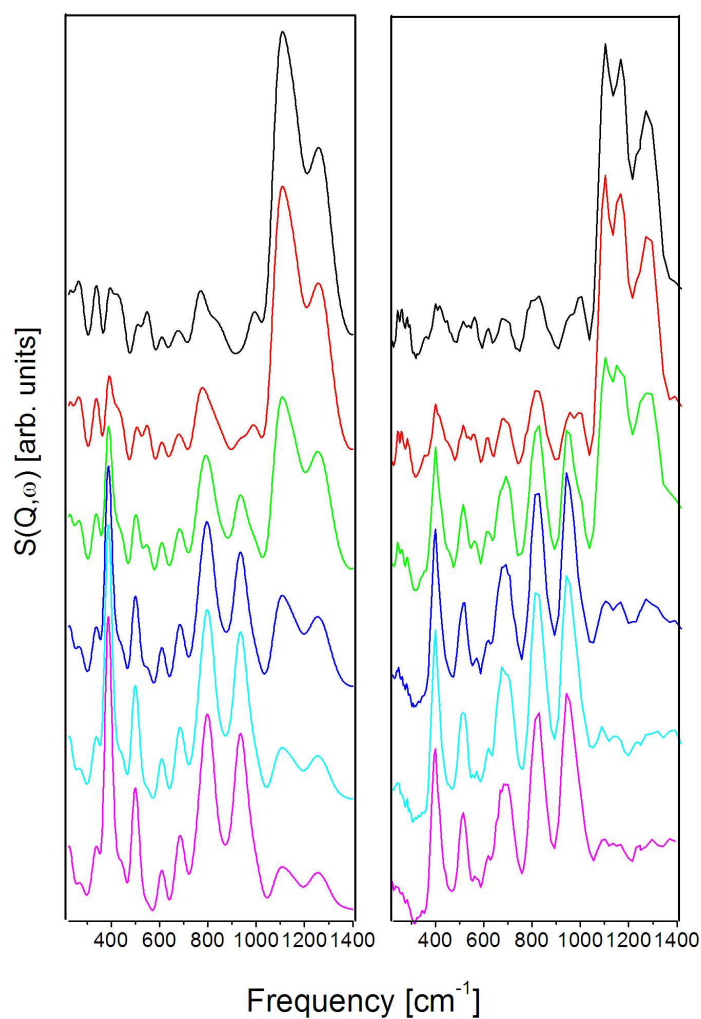

Figure 64

Comparison of measured (left) and calculated (right) INS vibrational spectra of PPTA for fibre orientations varying from parallel to the momentum transfer vector Q (top) to perpendicular (bottom).

D/5/ Discussion

The origin of the variations in the spectra of Figure 63 in terms of intra or inter-sheet interactions can be considered in terms of the structural characteristics of the different systems. The hydrogen bond geometry, described above for the Northolt structure, is

common to all systems. The polymer chain structure is characterised by the torsion angle between the amide linkage and the phenyl group and the bond angles ($\angle C_{phenyl}NC_{amide}$ and $\angle C_{phenyl}C_{amide}N$) within the amide linkage. To within one degree, the Northolt, *Pb* and Liu structures have the same charatceristic angles of 150°, 124° and 114°, thus they have the same hydrogen-bonded sheet structure. The parallel packing of the phenyl groups in the Northolt-translated structure, as opposed to herring-bone for the other structures, results in a torsion angle of 159°, the other angles being unchanged. In terms of the packing of amide linkages between hydrogen-bonded sheets, the Northolt and *Pb* systems with a parallel alignment, present a similar structure in which the two nitrogen-nitrogen distances between a amide linkage and the two nearest groups in a neighbouring sheet are almost identical (~4.7 Å). In the case of anti-parallel packing of amide linkages in the Northolt-translated and Liu systems, these two distances are different (4.3 Å and 5.0 Å). The Northolt and *Pb* systems therefore present almost identical 2D sheet structures and packing of phenyl groups and amide linkages between sheets and the spectra of vibrations are similar. The Liu system has a similar sheet structure but a different packing of amide linkages and the spectrum differs considerably from the Northolt and Pb spectra. Finally the Northolt-translated structure has a different sheet structure in terms of the phenyl-amide torsion but the same packing of amide linkages as in the Liu system and the spectrum ressembles that of the Liu structure. Clearly the spectra of vibrations are most sensitive to the inter-sheet interactions and to a lesser extent, the modification of the sheet structure (phenyl-amide torsion) by these interactions.

The VASP/DFT method can also be used to investigate the relative stability of the four structures. By comparing the total energies of the optimised structures in the same experimentally-determined unit cell, the Northolt, Liu and *Pb* structures, which all exhibit herring-bone packing of the phenyl groups between neighbouring sheets, are found to have similar stabilities (within 1 kcal/mol). The "Northolt-translated" structure, with its parallel packing of phenyl groups is more than 20 kcal/mol less stable than the other three structures. Therefore these calculations clearly exclude the parallel packing of phenyl groups in the solid state, but the effect of the relative orientation of amide linkages between sheets (parallel for Northolt and *Pb* structures, anti-parallel for Northolt-translated and Liu (*Pa*) structures) is more subtle. Assigning

reasonable partial charges to the four atoms in each of these three amide groups (q(O) = -0.60 e, q(C) = 0.45 e, q(N) = -0.49 e, q(H) = 0.28e) and calculating the Van der Waals (parameters from the Universal Force Field [114]) and Coulomb contributions to the total energy show that Coulomb interactions favour, by a few kcal/mol, the packing of amide linkages found in the anti-parallel case. However, this simple analysis of the inter-sheet packing of amide linkages depends on the partial charges and VDW parameterisation. While the packing of phenyl groups can be determined from the energies of the equilibrium structures, the packing of the amide groups gives much smaller energy differences and the role of vibrational spectroscopy coupled with DFT calculations in clarifying this structural issue is therefore crucial.

The structural information used as input to the phonon calculations was obtained at room temperature whereas the INS spectra were measured at 10 K. Using the spectra to identify the correct structure therefore depends on there being no major structural changes in cooling from room temperature to 10 K. That the calculation spectra based on the Liu structure are in good agreement with the measured spectra suggests that this structure is stable down to low temperatures.

Vibrational spectroscopy gives clear structural information inspite of disorder in the sample, arising from the imperfect crystallinity and the small fibre diameter, which limits spectral resolution. The sample-limited resolution is illustrated by the fact that the higher resolution TOSCA spectrometer did not give better resolved spectra than IN1. The phonon calculations on PPTA show weak dispersion, the width of the spectral contributions of the amide linkage hydrogen atom being smaller than 20 cm^{-1} (see for example the amide group hydrogen bond intensity at 800 cm^{-1} in Figure 63, panel b). This hydrogen atom participates in the hydrogen bond network and its modes could be expected to show dispersion, but the length of the hydrogen bond (the N-H…O distance is 2.98 Å) and the high N-H stretch frequency (~3400 cm^{-1}) indicate a weak hydrogen bond, which is consistent with the lack of significant dispersion. The ~30% of polymer chains on the outer surface of the fibres have vibrations modified by the molecular environment, but the remaining 70% of chains therefore have well-defined vibrations that are typical of the bulk material. If the coupling between molecules, indicated by more pronounced dispersion, was stronger, the effect of the surface would propagate further into the bulk and reduce the bulk signal. In this

context, the total vibrational spectrum has been compared with calculations performed on perfect, infinite crystals.

In a periodic model, surfaces can be modelled by removing some of the molecules, leaving a slab with an upper and a lower surface. Large models are therefore required to separate the two surfaces both via the vaccum and via the bulk. The calculations should be repeated for different molecular planes at the surface. Since good agreement has been obtained between calculations perfomed on the bulk material and measurements on fibres, it is expected that, in the case of PPTA, the effect of the fibre surfaces is mainly to broaden a fraction (~30%) of the bulk signal. Separating quantitatively, the crystalline and amorphous contributions are therefore difficult. However, quantifying different crystalline contributions in a sample would be possible, with a precision dependent on there being distinct spectral features associated with each component.

D/6/ Conclusion

The calculations on benzanilide demonstrate how well solid state DFT methods describe intra and intermolecular interactions. Applying the same method to PPTA gives clear evidence for an anti-parallel configuration of amide linkages in neighbouring hydrogen-bonded sheets. The calculations presented here are therefore in agreement with the *Pa* structure determined by Liu [44] and the new neutron diffraction data [42]. It is the accurate calculation of the weak inter-sheet interactions, in the presence of stronger intra-sheet interactions, that enables the structures to be distinguished. The N-H wagging mode is the principle probe of the inter-sheet packing since the hydrogen atom can have a large amplitude motion (when this mode is pure) giving an intense peak.

In DFT-based calculations, the total energy is not determined from a pair-wise sum of inter-atomic interactions so the contributions from different structural units cannot be determined by "tuning" empirical atom-atom potentials. It is only by repeating the phonon calculations on a set of structures that the key structural feature probed by the vibrations, the relative orientation of the amide linkages in neighbouring hydrogen-bonded sheets, can be identified. While total energy calculations can exclude the

Northolt-translated structure due to its parallel packing of phenyl groups, the different packing of amide groups gives much smaller energy differences and the vibrational spectroscopy therefore provides uniquely clear information.

Since INS intensities can be calculated reliably from a given structure and its vibrational modes, the validity of the structure can be established even when the spectral resolution is limited and individual modes are not resolved. Thus by using a fully converged, first principles method that allows the whole spectral profile (peak positions, intensities and shapes) to be calculated, a direct link is established between the structure and dynamics of these model amide compounds. In view of the small shifts of vibrational modes that change the overall spectral profile, any parametrised description of the interatomic interactions could not be relied upon to obtain such precise structural information.

As computing power continues to grow and with the development of more efficient computational codes, the methods used here will be extended to more complex systems of biological and technological interest. Intra-sheet modes in β-sheet compounds could be investigated, as has been done elsewhere [105,106], but avoiding the transfer of parameters from *ab initio* calculations to empirical potentials. Assemblies of amino acids or nucleic acids in their natural aqueous environment, such as collagen, small globular proteins and DNA, are also coming within the scope of these methods. The partially crystalline nature of such systems results in a relatively high level of spectral information, justifying the use of computationally intensive, first principles methods. However, due to the presence of non-crystalline water, a molecular dynamics approach, in which a range of conformational states of the system is explored and the vibrational spectrum is calculated from the velocity auto-correlation function, may be preferred to the structure-specific, dynamical matrix approach employed here. As the work presented here demonstrates, interest need not be limited to particular vibrational bands, the whole spectral profile and its information content can be investigated. By calculating charge tensors, the calculation of spectral profiles can also be extended from INS spectra to the more commonly measured IR spectra.

Appendix E

A spinning system for the study of fibrous samples by elastic and inelastic neutron scattering

E/1/ Manual for the spinning apparatus [115]

E/1/1/ Description of the setup:

The machine is made of the DNA solution seringe, the spinneret, the precipitation column, the spinning bath and the degazor.

E/1/2/ DNA solution preparation

Dissolve the DNA into a solvent made of 0.15M NaCl, 0.003M Na Citrate, 5e-4M Na-EDTA.

Homogeneise the solution

Adjust the optical density of the solution to 24 at 260nm (roughly 1mg/ml)

Heat up the solution to room temperature, and degase it for a few minutes.

Clean the spinneret with demineralised water and checking that the wholes are not blocked. Suck demineralised water into the spinneret + small teflon tubing (check that no air bubble is stucked) and connect the tubing to the seringe. Install the seringe at its place on the apparatus, and keep the spinneret into a bath of water.

E/1/3/ Column preparation

Place the cylinder at initial position, adjust the cogwheels and the back-and-forth switches. Place the small plate of the desired width between the 2 switches. Check the length : 1.25 mm per turn between the two 'clics'. Adjust the column at the right place vertically (about 5mm up to the cylinder, higher makes the fibres more spread over the cylinder, lower is good for thin films) and horizontally (3/4 of the radius).

Fill the flak connected to the ethanol pump with pure ethanol.

E/1/4/ Spin bath preparation

Prepare 6 to 7 litres of the spinning bath made of 75% Ethanol and <0.25M NaCl. Degase the bath into the intermediate flask. Fill the cuve with the bath, fill the degasor in actioning the circulation pump, and refill the cuve up to the top.

Place the spinneret into the column, and suck the bath up to the top of the column.

Switch on the rotational motor and up the cylinder of air bubbles with a glass tool. Place the 'trying' area of the cyclinder under the guide.

E/1/5/ Spinning the film.

Switch on the motor of the seringe and the circulation pump. Check that the ethanol pump is working, the circulation pump too, and that liquid comes out of the spinneret. Take the Allan's position and wait for the fibres to come! Grab the fibre inside the guide and take it around the cylinder. When spinning is regular, place the cylinder into the spinning area (quickly turn the wheel until the first clic), start the axial motion and the clock. Start the spinning for the number of layers you want.

At the end of the last layer, quickly comes out of the spinning area and stop the axial motion, the circulation pump and the seringe motor.

Clean up the undesired fibres, empty the cuve by sucking the bath until only the last part of the cylinder soaks into the bath. Stop the rotational motion and remove the cylinder, keep it into a bath.

E/1/6/ Crystallisation and drying procedure.

Put the cylinder into a bath made of 77 or 75% EtOH, 0.03M NaCl and keep it with a magnet steerer over night at 5 C. The concentration of ethanol and salt of this solution can be adjusted for optimal crystallisation : the NaCl concentration should be less than < 0.01 M, or another salt (LiCl, CaCl2, MgCl2, RbCl2…) can be used at this point. In that case, use several bath (3 to 4) for total exchange.

Take the cylinder out of the bath, place it on a rotating stick and absorb the bath around the film. Place the cylinder into a cold dissecator (that can contains 0.5 g silica

gel or not, depending on the crystallisation speed) until it is dry : no more sticky under the fingers. Place a NaCl saturated solution into the dissecator, and take it slowly to room temperature in keeping it into a polystyrene box.

E/1/7/ Cut and fold

Check that the film looks a bit silvery. Take the cylinder out, wait a minute at room temperature and humidity, and cut the film at once with a razorblade. Help it gently taking off the cylinder with a spatule, and grab it with handing tweezers. The film contracts at room humidity (switches from the B from on the cylinder to the A form at 50% humidity).

E/1/8/ LiDNA and other salts.

LiDNA can be made using LiCl into the spinning bath and all other baths. Ethanol concentration is higher (80%). Ca, Mg… DNA are made by ion exchange from NaDNA. In that case, check the ethanol concentration of the baths (72%…).

Comments

Quality of the film depends on the purity of DNA, length of the strands (highly polymerised). Ethanol has to be ultra pure (99.5%)

E/1/8/ Pictures of the spinning apparatus
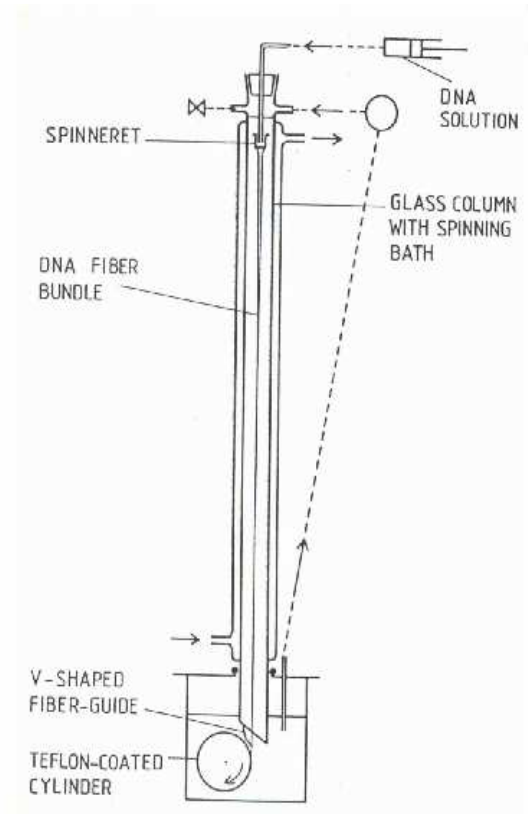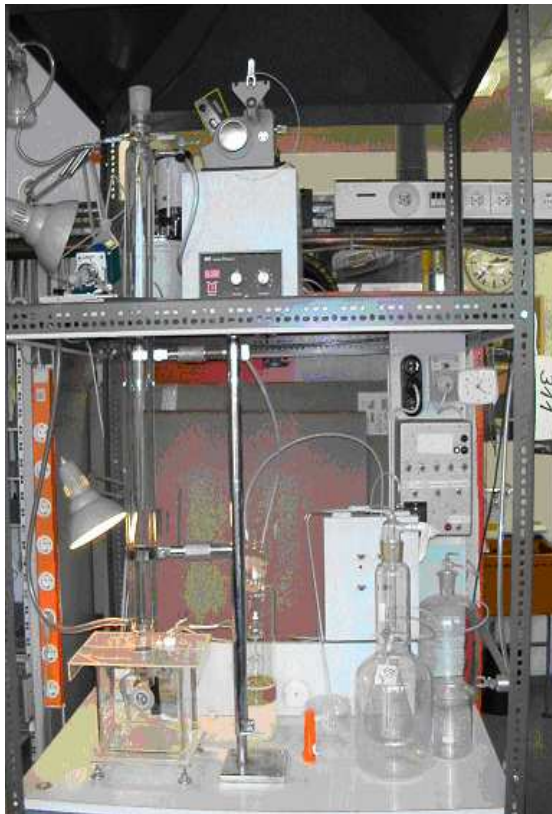
Picture 1

Left: spinning apparatus, right: different elements of the machine

Picture 2 left and right

Spinning the DNA film around cylinder
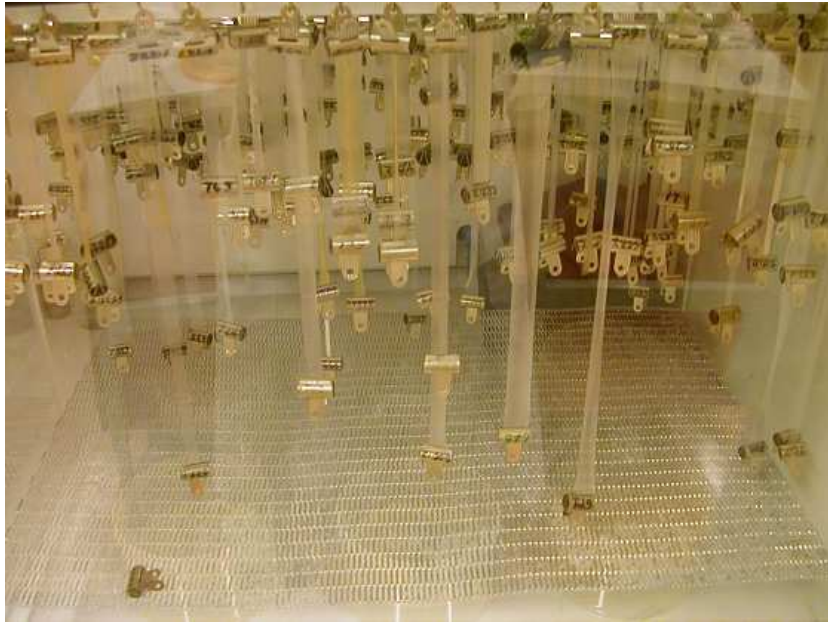
Picture 3

Films after cutting and drying procedure

Appendix F

DNA

## F/1/ Conformational variants of the DNA double helix

B-DNA	A-DNA	C-DNA	D-DNA	Z-DNA
right-handed	right-handed	right-handed	right-handed	left-handed
10 base pairs per turn	11 base pairs per turn	9.3 base pairs per turn	8 base pairs per turn	12 base pairs per turn
pitch = 34 Å	pitch = 28.2 Å	pitch = 31 Å	pitch = 24.2 Å	pitch = 43 Å
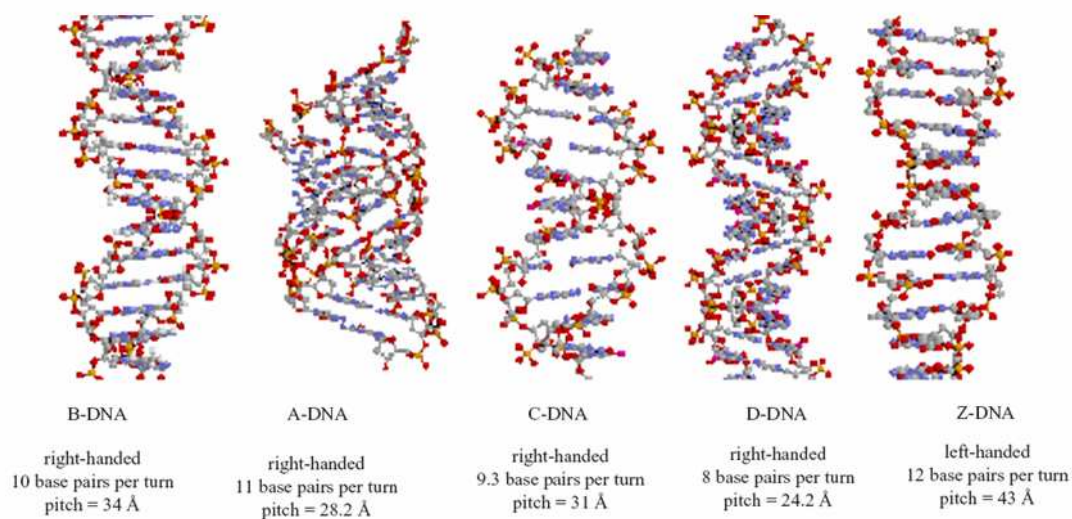
F/2/ Water sites in A-DNA [116]
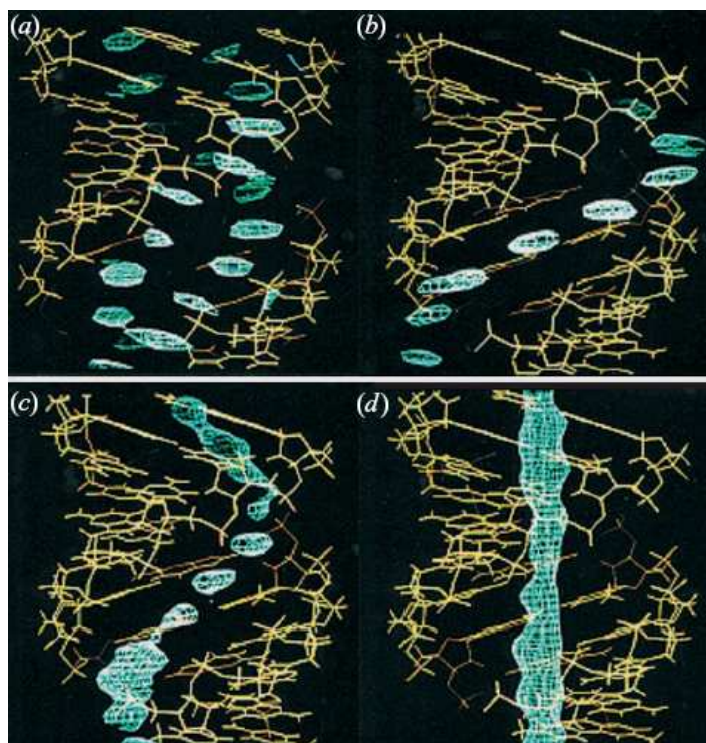
(a) Ordered water along the edges of the major groove linking successive O1 phosphate oxygen atoms, (b) ordered water running along the outside of the major

groove, (c) a string of water features located deeper in the centre of the groove, (d) a central column of water running along the axis of the double helix. The interpretation of the last of these features is complicated by the sequence averaging that occurs in x-ray diffraction experiments on this "mixed-sequence" DNA.

F/3/ Vibrations

$\omega = 228cm^{-1}$, mode number 2360: methyl rotation of T (in light green)
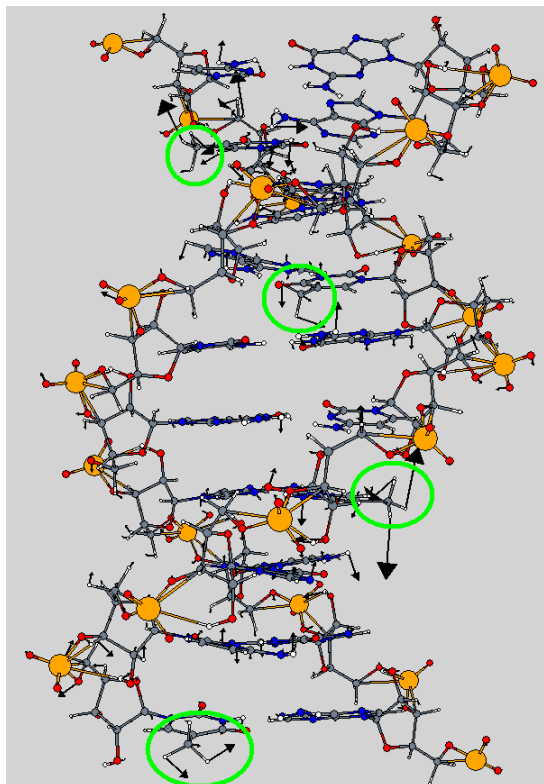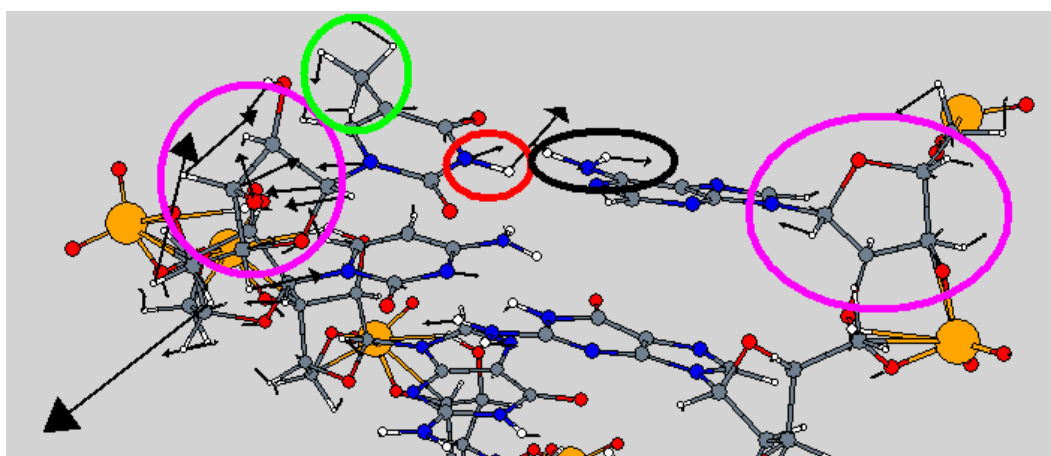

$\omega = 407cm^{-1}$, num. 4585: op NH2 wag of G (in green) and C (in blue)
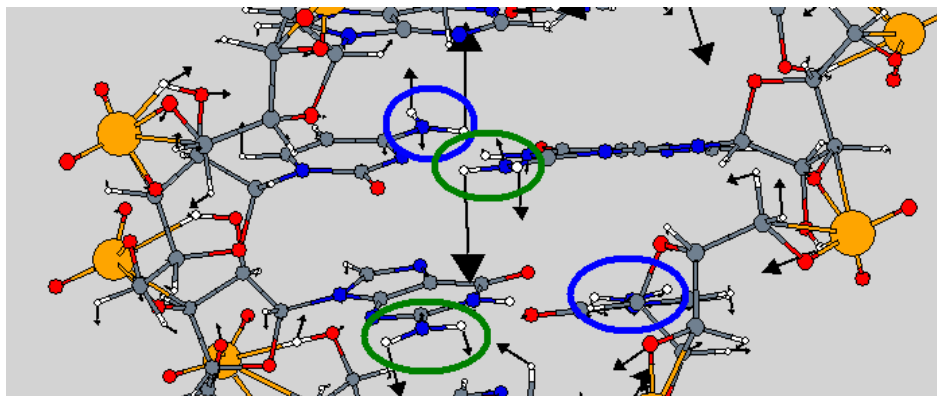
ω = 520cm-1, num. 5713: op NH2 wag of A (black), G (green), C (blue) + op NH wag of T (red)
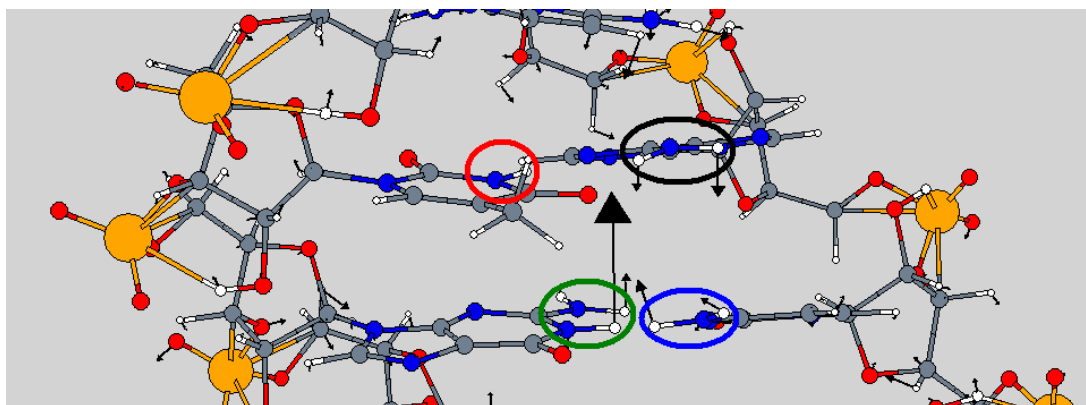
ω = 571cm-1, num. 6036: op NH2 wag of A (black) and C (blue) + op NH wag of T (red) + sugar (pink) and Phosphate group deformations
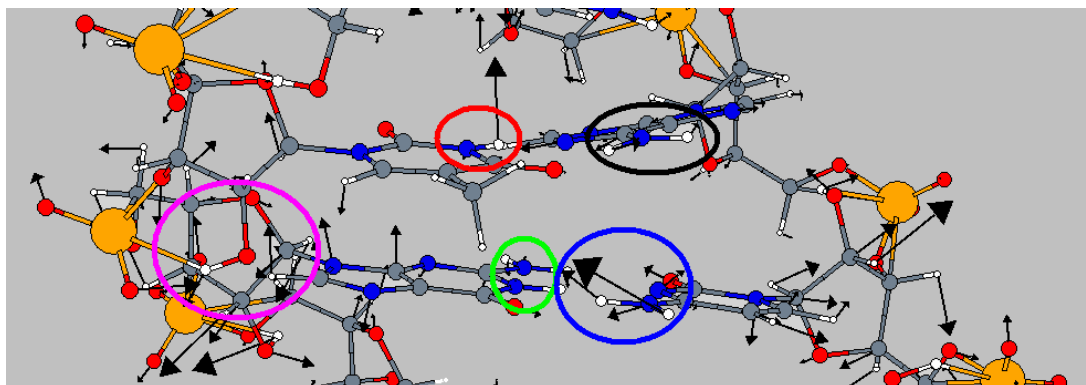
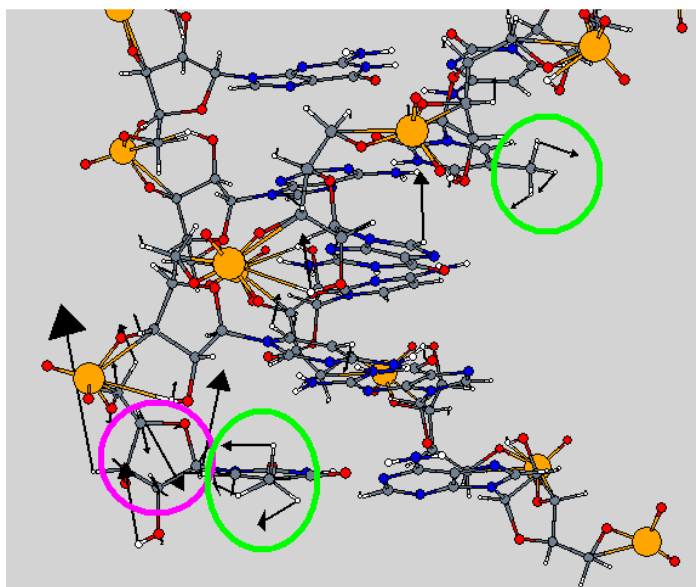ω = 807cm-1, num. 6717: ip NH2 wag of A (black), CH3 deformation (light green) and NH wag of T (red)+ sugar deformations (pink)
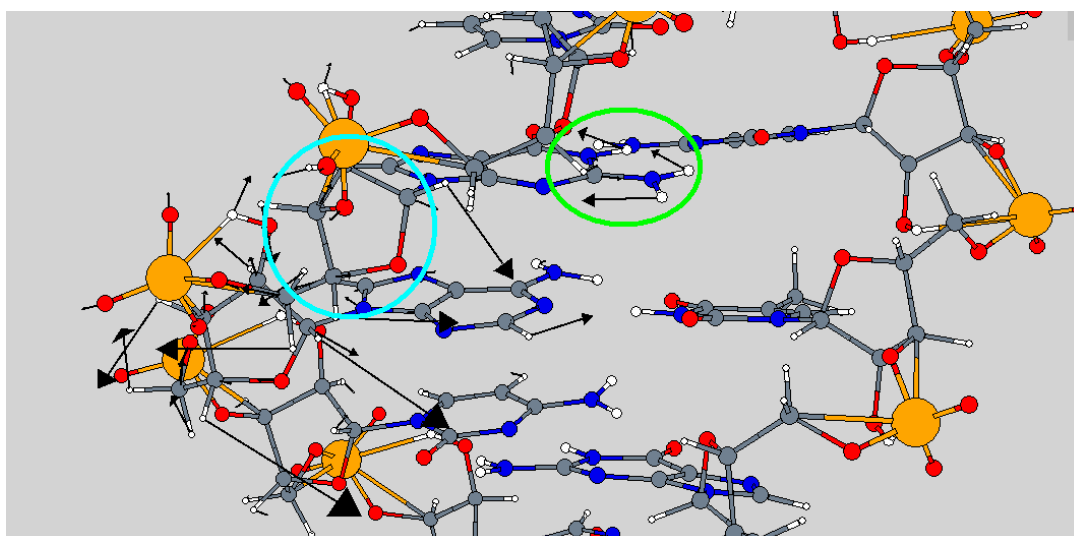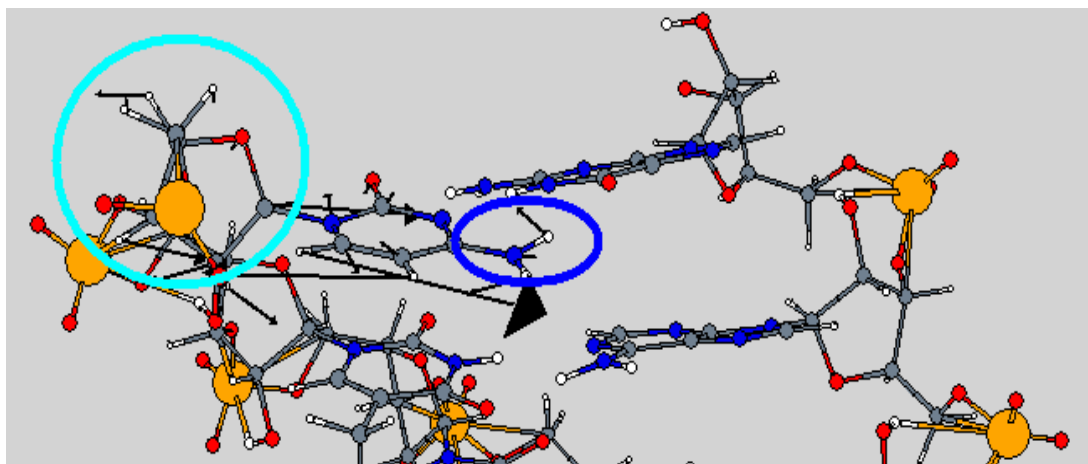
ω = 913 cm-1, num. 6810: sugar deformations
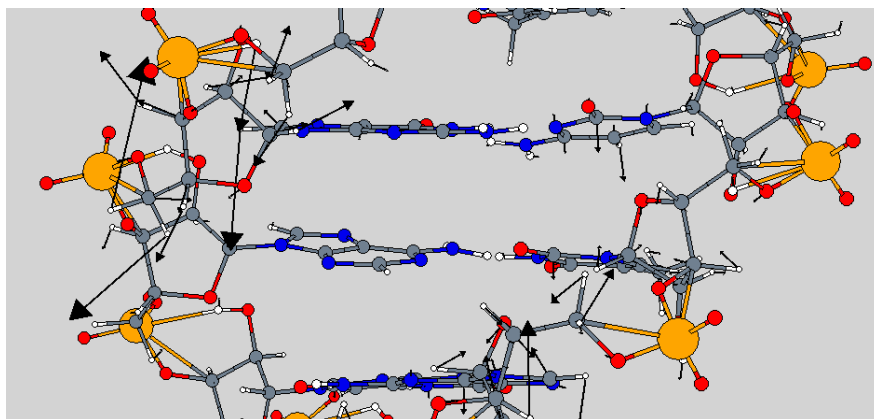
ω = 1009 cm-1, 6910: sugar (pink) +CH3 deformation of T (light green)

ω = 1109 cm-1, 7001: sugar + ip NH2 wag of G (green)

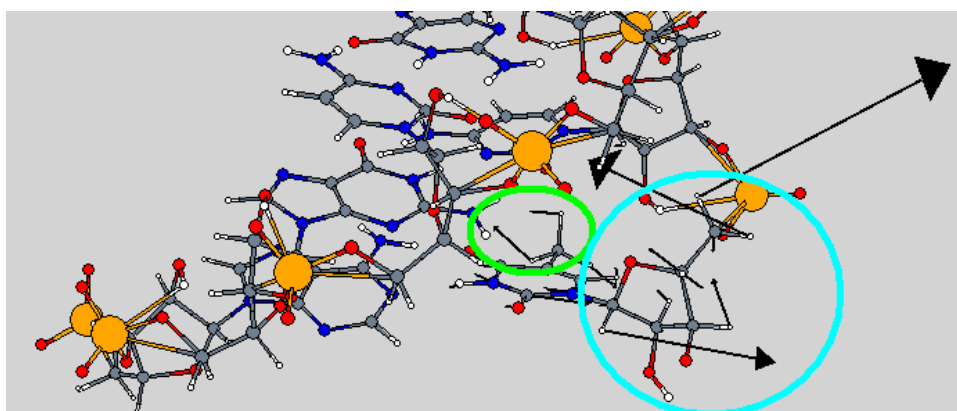ω = 1257cm-1, num. 7162: ip NH2 bend of C (blue) +sugar deformations (turquoise)

ω = 1385cm-1, num. 8265: sugar
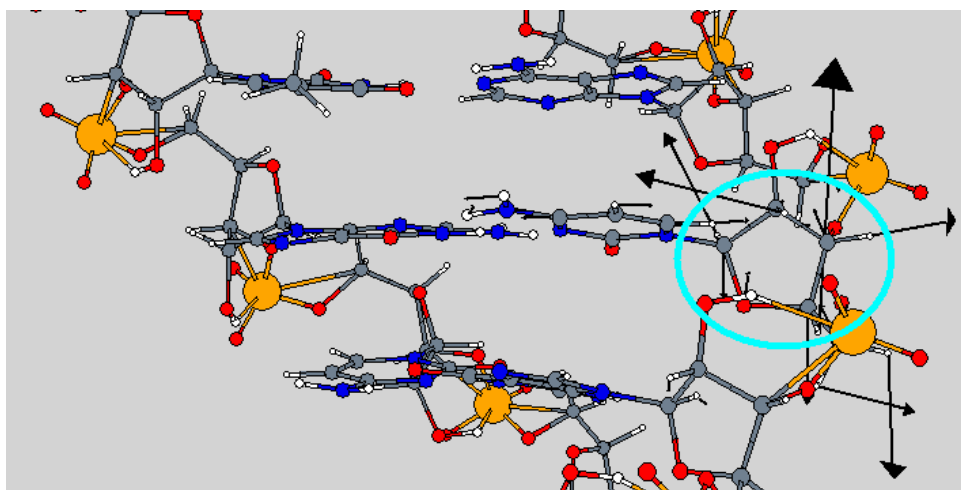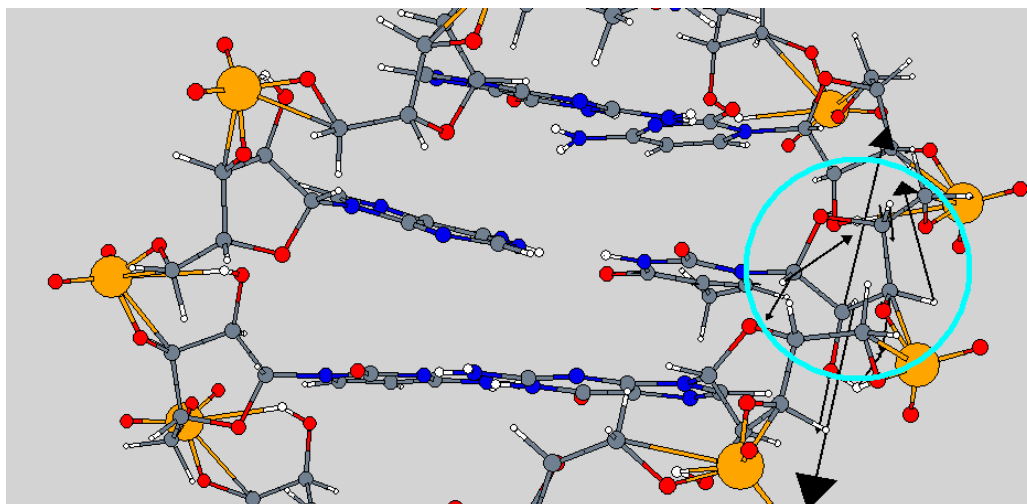

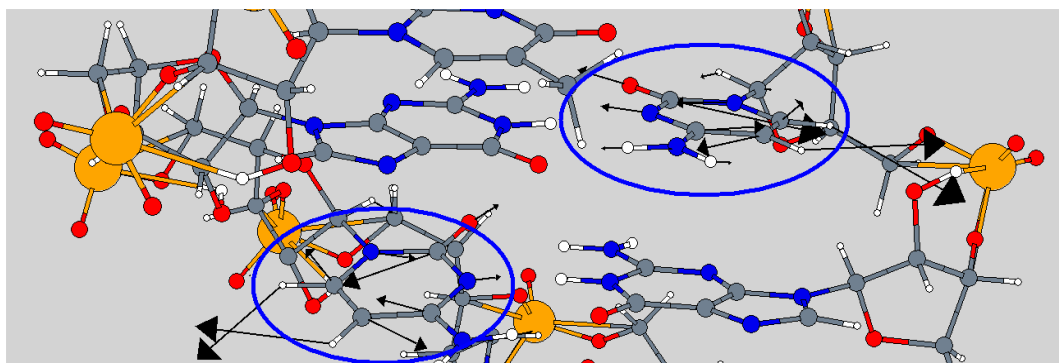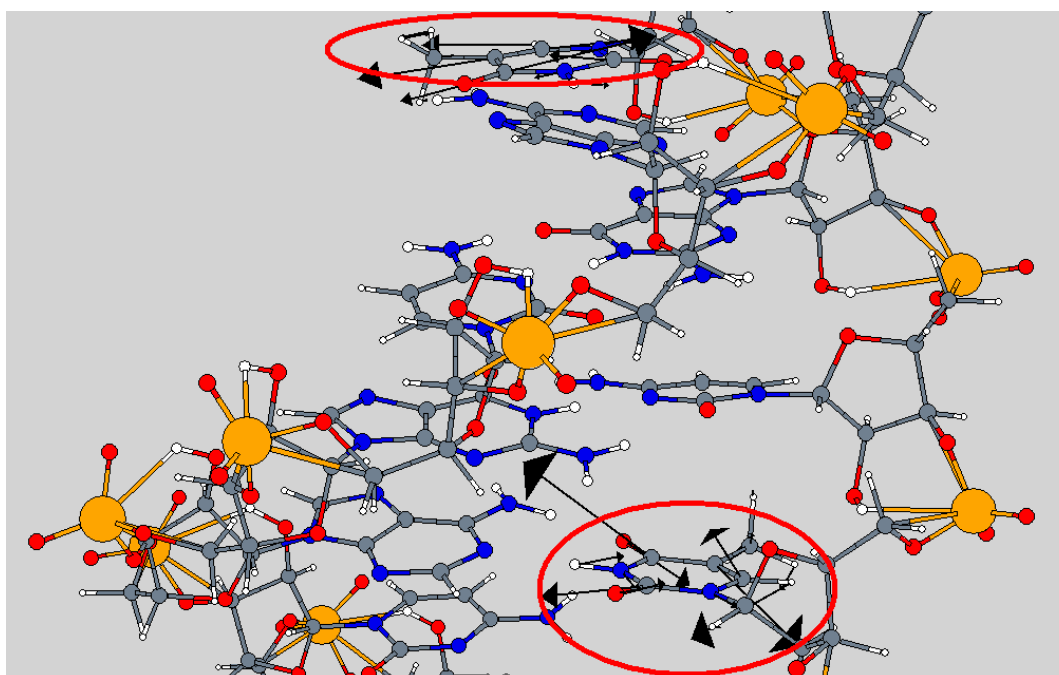ω = 1435cm-1, 8315: sugar (turquoise) +CH3 deformations of T (light green)

ω = 1638cm-1, 8458: sugar deformations

ω = 1725cm-1, num. 8490: C

ω = 1821cm-1, num. 8505: T

References

[1] Fillaux F., Nicolaï B., Chem. Phys. Lett., 2005, 415, 357-361

[2] Jeffrey G.A., 1997, an Introduction to Hydrogen Bonding, Oxford University Press

[3] Kearley, G.J.; Johnson, M.R.; Plazanet, M.; Suard, E.; J. Chem. Phys., 2001, 115, 2614-2620

[4] Plazanet, M.; Fukushima, N.; Johnson, M.R.; Horsewill. A.J.; Trommsdorff, H-P.; J. Chem. Phys. 2001, 115, 3241

[5] Plazanet, M.; Fukushima, N.; Johnson, M.R.; Chem. Phys., 2002, 280, 53-70

[6] Johnson, M.R.; Trommsdorff, H-P.; Chem. Phys. Lett., 2002, 364, 34-38

[7] Johnson, M.R.; Parlinski, K.; Natkaniec, I.; Hudson, B.; Chem. Phys., 2003, 291, 53-60

[8] Neumann M., Brougham D.F., McGloin C.J., Johnson M.R., Horsewill A.J., Trommsdorff H.P., J. Chem. Phys., 1998, 109, 7300-7311

[9] Watson J.D., Crick F.H.C., Nature, 1953, 171, 737-738

[10] Hübsch A., Endres R.G., Cox D.L., Singh R.R.P., Phys. Rev. Lett., 2005, 94, 178102

[11] Gervasio L., Carloni P., Parrinello M., Phys. Rev. Lett., 2002, 89, 108102

[12] Cleland W.W., Kreevoy M.M., Science 1994, 264, 1887-1890

[13] Kim K.S., Oh K.S., Lee J.Y., Proc. Natl. Acad. Sci., 2000, 97, 6373-6378

[14] Pauling L., Corey R.B., J. Am. Chem. Soc., 1950, 72, 5349

[15] J. A. Cowan, thesis, 2002

[16] J. A. Cowan, J.A.K. Howard, J. McIntyre, S.M.-F. Lo and I.D. Williams, Acta.Crystallogr. B61, 724 (2005)

[17] Reiter G.F., Mayers J., Platzman P., Phys. Rev. Lett., 2002, 89, 135505

[18] Fisher S., Smith J.C., Verma C.S., J. Phys. Chem. B, 2001, 105, 8050-8055

[19] Zhou H.X., Proc. Natl. Acad. Sci., 1998, 95, 9280-9283

[20] Frauenfelder H., Slinger S.G., Wolynes P.G., Science, 1991, 254, 1598

[21] Frauenfelder H., Parak F., Young R.D., Annu. Rev. Biophys. Biophys. Chem., 1988, 17, 451

[22] Moritsugu K., Miyashita O., Kidera A., Phys. Rev. Lett., 2000, 85, 3970-3973

[23] Moritsugu, J. Phys. Chem. B, 2003, 107, 3309-3317

[24] Franks F., Biophysical Chemistry, 2002, 96, 117-127

[25] Ferrand M., Dianoux A.J., Petry W., Zaccai G., Proc. Natl. Acad. Sci., 90, 9668-9672

[26] Réat V., Dunn R., Ferrand M., Finney J.L., Daniel R.M., Smith J.C., Proc. Natl. Acad. Sci., 2000, 97, 9961-9966

[27] Paciaroni A., Cinelli S., Cornicchi E., De Francesco A., Onori G., Chem. Phys. Letters, 2005, 410, 400-403

[28] Svergun D.I., Richard S., Koch M.H.J., Sayers Z., Kuprin S., Zaccai G., Proc. Natl. Acad. Sci., 1998, 95, 2267-2272

[29] Tarek M., Tobias D.J., Biophysical J., 2000, 79, 3244-3257

[30] Russo D., Baglioni P., Peroni E., Teixeira J., Chem. Phys., 2003, 292, 235-245

[31] Elliott A., Ambrose E., Nature, 1950, 165, 921-922

[32] Bamford C.H., Elliot A., Hanby W.E., Synthetic Polypeptides (Academic Press, New York), 1956

[33] Kneipp, J.; Miller, L.M.; Joncic, M.; Kittel, M.; Lasch, P.; Naumann, D.; Biochim. Biophys. Acta. 2003, 1639, 152

[34] McColl, I.H.; Blanch, E.W.; Gill, A.C.; Rhie, A.G.O.; Ritchie, M.A.; Hecht. L.; Nielsen, K.; Barron, L.D.; J. Am. Chem. Soc. 2003, 125, 10019

[35] Prusiner S.B., Science 1997, 278, 245

[36] Krimm S., Abe Y., Proc. Natl. Acad. Sci., 1972, 69, 2788-2792

[37] Krimm S., Bandekar J., Adv. Protein Chem., 1986, 38, 181-364

[38] Shifu W., Roberts J., Kempfert K., Nicolet Application Note, 8

[39] H.D.Middendorf, R.L.Hayward, S.F.Parker, J.Bradshaw, and A.Miller, Biophysical J. (1995), vol.69, 660-673

[40] Northolt, M.G.; van Arsten, J.J.; J. Polym. Sci. Polym. Lett. Ed. 1973 11 333

[41] Northolt, M.G.; Europ. Polym. J. 1974, 10, 799

[42] Gardner, K.H.; English, A.D.; Forsyth, V.T.; Macromolecules

[43] Tashiro, K.; Kobayashi, M.; Tadokoro, H.; Macromolecules, 1977 10, 413

[44] Liu. J.; Cheng, S.Z.D.; Geil, P.H.; Polymer 1996, 8, 1413

[45] Lotz B., J. of Mol. Biol., 1974, 87, 169

[46] F.Crick and A.Rich, Nature (1955), vol.176, 780-781

[47] K.Taga, M.G.Sowa, J.Wang, H.Etori, T.Yoshida, H.Okabayashi, H.H.Mantsch, Vibrational Spectroscopy, 14 (1997), 143-146

[48] Litovchenko G.D., Sokolova T.S., Volokhina A.V., Kudryavtsev G.I., Papkov S.P., Journal of Applied Spectroscopy, 20, 3, 345-348

[49] S.Arnott and S.D.Dover, Acta Cryst.(1968), B24, 599-601

[50] Gupta V.D., Singh R.D., Dwivedi A.M., Biopolymers, 1973, 12, 1377-1385

[51] Sakakibara, S.; Kishida, Y.; Kikuchi, Y.; Sakai, Y.; Kakiuchi, K., 1968, Bull.Chem.Soc.Japan, 41, 1273

[52] G.N.Ramachandran and R.Chandrasekharan, Biopolymers (1968), vol.6, 1649-1658

[53] Diem M., Bhatnagar R.S., Druyan M.E., Renugopalakrishnan V., Biopolymers, 1984, 23, 2955-2961

[54] Lazarev Yu A., Biopolymers, 1985, 24, 1449-1478

[55] Berney C.V., Biophysical J., 1987, 52, 343-345

[56] M. Peyrard and A. R. Bishop, Phys. Rev. Lett. 62 (1989) 2755

[57] T. Dauxois et al., Phys. Rev. E 47, (1993) 684

[58] T. S. van Erp et al., Phys. Rev. Lett., 95 (2005) 218104

[59] Urabe H., Hayashi H., Tominaga Y., Nishimura Y., Kubota K., Tsuboi M., J. Chem. Phys., 82, 531-535

[60] Tominaga Y., Shida M., Kubota K., Urabe H., Nishimura Y., Tsuboi M., J. Chem. Phys., 83, 5972-5975

[61] Grimm H., Stiller H., Majkrzak C.F., Rupprecht A., Dahlborg U., Phys. Rev. Lett., 1987, 59, 1780-1783

[62] Liu Y., Berti D., Faraone A., Chen W.R., Alatas A., Sinn H., Alp E., Said A., Baglioni P., Chen S.H., Phys. Chem. Chem. Phys., 2004, 6, 1499-1505

[63] Liu Y., Chen S.H., Berti D., Baglioni P., Alatas A., Sinn H., Alp E., Said A., JCP, 2005, 123, 214909

[64] Prohofsky, E. W., Phys. Rev. A 38, 1988, 1538–1541

[65] Chen, Y. Z., and E. W. Prohofsky. 1995, Biopolymers. 35:657–666.

[66] Hua, X. M., and E. W. Prohofsky. 1988, Biopolymers. 27: 645–655.

[67] Peyrard M., Bishop A.R., Phys. Rev. Lett., 1988, 62, 2755-2758

[68] Voulgarakis N.K., Kalosakas G., Rasmussen K. Ø., Bishop A.R., nanoletters 4, 629

[69] Parlinski K.; Am.Inst.Phys. Conference Proceedings 479," Neutrons and numerical methods N2M", edited by M.R.Johnson, G.J.Kearley, and H.G.Buttner, 1999, 121-130

[70] Plazanet M., Fukushima N., Johnson M.R., Chem Phys 280, 53-70

[71] Ambjörnsson T., Metzeler R., PACS, 87, 1-4

[72] Koch W., Holthausen M.C., A Chemist's Guide to Density Functional Theory, 2nd Edition, 2001

[73] P. Hohenberg and W. Kohn, Phys. Rev., 1964, 136, B864-B871

[74] W. Kohn and L. J. Sham, Phys. Rev., 1965, 140, A1133

[75] http://www.isis.rl.ac.uk/molecularspectroscopy/tosca/

[76] http://www.ill.fr/YellowBook/IN1/

[77] http://www.ill.fr/YellowBook/IN4/

[78] http://www.ill.fr/YellowBook/IN5/

[79] Katz, B.A.; Spencer, J.R.; Elrod, K.; Luong, C.; Mackman, R.L.; Rice, M.; Sprengeler P.A.; Allen, D.; Janc, J.; J. Am. Chem. Soc. 2002, 124, 11657-11668

[80] Anderson , S.; Crosson, S.; Moffat, K.; Acta.Cryst. 2004, D60, 1008-1016

[81] Jones, P.A.; Takai, D. ; Science 2001 , Aug 10, 1068-1070

[82] Villani, G.; Chem. Phys. 2005, 316, 1-8

[83] Wilson, C.C.; Acta.Cryst. 2001, B57, 435-439

[84] Wilson, C.C.; Morrison, C.A.; Chem.Phys.Lett. 2002, 362, 85-89

[85] Morrison, C.A.; Siddick, M.M.; Camp, P.J.; Wilson, C.C.; J. Am. Chem. Soc 2005, 127, 4042-4048

[86] Wilson, C.C.; Shankland, N.; Florence, A.J.; J Chem Soc Faraday Trans 1996 92, 5051-5057

[87] Rog, T.; Murzyn, K.; Hinsen, K.; Kneller, G.; J. Comp. Chem. 2003, 24, 657

[88] http://cms.mpi.univie.ac.at/vasp/

[89] Kresse, G.; Furthmuller, J.; Phys.Rev. 1996, B54, 11169-11186

[90] Kresse, G.; Furthmuller, J.; Comput.Mat.Science 1996, 6, 15-50

[91] Kresse, G.; J.Hafner, Phys.Rev. 1993, B47, 558-561

[92] Perdew, J.P.; Wang, Y.; Phys. Rev. 1992, B 45, 13244-13249

[93] Vanderbilt, D.; Phys.Rev. 1990, B41, 7892-7895

[94] Rodrigues, B.L.; Tellgren, R.; Fernandes, N.G.; Acta.Cryst. 2001, B57, 353-358

[95] Gilli, P.; Bertolasi, V.; Ferreti, V.; Gilli, G.; J. Am. Chem. Soc., 1994, 116, 909

143

[96] Ilczyszyn, M.M.; Ratajczak, H.; Barnes, A.J.; J. Raman Spectroscopy, 1992, 23, 1-13

[97] Luck, W.; Wess, T.; J. Mol. Struct. 1992, 270, 229-245

[98] J. A. Cowan, J.A.K. Howard, G.J. McIntyre, S. Lo and I.D. Williams, Acta.Crystallogr. B59, 794 (2003)

[99] T. Steiner, C.C. Wilson and I. Majerz, Chem. Commun. 14, 1231 (2000)

[100] F. Fontaine-Vive, M.R. Johnson, G.J. Kearley, J.A.K. Howard and S.F. Parker, J. Am. Chem. Soc. 128, 2963 (2006)

[101] M. Plazanet, F. Fontaine-Vive, K.H. Gardner, V.T. Forsyth, A. Ivanov, A.J. Ramirez-Cuesta and M.R. Johnson, J. Am. Chem. Soc. 127, 6672 (2005)

[102] M. Neumann, D.F. Brougham, C.J. McGloin, M.R. Johnson, A.J. Horsewill and H-P. Trommsdorff, J. Chem. Phys., 109, 7300 (1997)

[103] Kneipp, J.; Miller, L.M.; Joncic, M.; Kittel, M.; Lasch, P.; Naumann, D.; Biochim. Biophys. Acta. 2003, 1639, 152-158

[104] McColl, I.H.; Blanch, E.W.; Gill, A.C.; Rhie, A.G.O.; Ritchie, M.A.; Hecht. L.; Nielsen, K.; Barron, L.D.; J. Am. Chem. Soc. 2003, 125, 10019-10026

[105] Kubelka, J.; Keiderling, T.A.; J. Am. Chem. Soc. 2001, 123, 6142-6150

[106] Kubelka, J.; Keiderling, T.A.; J. Am. Chem. Soc. 2001, 123, 12048-12058

[107] Yang, H.H., "Kevlar Aramid Fiber", Wiley, New York, 1993

[108] Jackson, C.L.; Schadt, R.J.; Gardner, K.H.; Chase, D.B.; Allen, S.R., Gabara; V.; English, A.D.; Polymer 1994, 35, 1123

[109] Shen, D.Y., Molis, S.E. and Hsu, S.L., Polymer Eng. And Sci. 1983, 23, 543-547

[110] Kim, P.K., Chang, C. and Hsu, S.L. Polymer 1986, 27, 34-46

[111] Rutledge, G.C.; Suter, U.W.; Papaspyrides, C.D. Macromolecules 1991, 24, 1934-1943

[112] Kim, P.K.; Hsu, S.L.; Ishida, H. Macromolecules 1985 18, 1905-1914

[113] Kashino, S.; Ito, K.; Haisa, M.; Bul. Chem. Soc. Japan, 1979, 52, 365-369

[114] Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. , J. Am. Chem. Soc., 1992, 114, 10024-10035

[115] A.Rupprecht, Acta Chemica Scandinavica, 20 (1966) 494

[116] Shotton et al., Biophys. Chem., 1997, 69, 85-96

Acknowledgment

Outside my thesis, I would like to express my gratitude to my family and friends who allowed me to forget my thesis. I also thank them for trying to understand and appreciate my thesis work.

I thank both of my supervisors for their helpful remarks and their patience.

I finally thank all the people who have contributed directly or indirectly to my thesis, who helped me of their own free will, or otherwise.

Curriculum vitae

I was born on the 15th of May in 1979. From 1994 to 1997, I attended the lycée Berthollet in Annecy. In 1997 I started my study in the engineering school of Toulon finishing in 1999. Then, I followed physical science courses at the University of Savoie in Chambéry and obtained in 2003 my master degree in chemical-physics at the University Joseph Fourier in Grenoble. Next, I accepted a PhD position shared between the Interfaculty Reactor Institute of the Delft University of Technology and the Institut Laue-Langevin in Grenoble. In this thesis, under the supervision of Prof. G. J. Kearley from Delft and Prof. M. R. Johnson from Grenoble, I describe my work on the dynamics of bio-molecular systems.