Chemical Physics 355 (2009) 141–148

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Collagen and component polypeptides: Low frequency and amide vibrations

F. Fontaine-Vive [a,b], F. Merzel [c], M.R. Johnson [a,*], G.J. Kearley [d]

[a] Institut Laue Langevin, Boîte Postale 156, 38042 Grenoble Cedex 9, France
[b] Radiation, Reactors and Radionuclides Department, Faculty of Applied Sciences, Delft University of Technology, Mekelweg 15, 2629 JB Delft, The Netherlands
[c] National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia
[d] Bragg Institute, Building 87, Australian Nuclear Science and Technology Organisation, PMB 1 Menai, NSW 2234, Australia

ARTICLE INFO

Article history:
Received 6 October 2008
Accepted 2 December 2008
Available online 10 December 2008

Keywords:
DFT
Collagen
Amide bands
Phonons
Breathing modes
Water dynamics

ABSTRACT

Collagen is a fibrous protein, which exists widely in the human body. The biomechanical properties of collagen depend on its triple helix structure and the corresponding low frequency vibrations. We use first-principles, density functional theory methods and analytical force fields to investigate the molecular vibrations of a model collagen compound, the results being validated by comparison with published, inelastic neutron scattering data. The results from these atomistic simulations are used at higher frequency to study the Amide I and V vibrations and therefore the vibrational signature of secondary and tertiary structure formation. In addition to collagen, its component homopolymers, poly-glycine and poly-proline are also studied. The Amide V vibration of glycine is strongly modified in going from the single helix of poly-glycine II to the triple helix of collagen. The collagen models are hydrated and this work allows us to discuss the relative merits of density functional theory and force field methods when tackling complex, partially crystalline systems.

1. Introduction

In biology, the hydrogen-bond is a corner stone of our understanding of living materials [1,2]. Intermediate in strength between covalent bonds and inter-molecular interactions, they can confer stability on molecular assemblies such as proteins and DNA and yet allow these structures to dissociate or denature under specific conditions. Moreover, these bio-molecular systems are solvated by water molecules which form extended hydrogen-bond networks.

Detailed understanding of the role of hydrogen-bonding in complex systems is difficult since crystallography gives accurate information on systems with long-range order with a precision that is inversely proportional to the size of the system. And yet a change in hydrogen-bond length by less than 0.1 Å can cause a pronounced change in the behaviour of the bond. Vibrational spectroscopy is a local probe of structure since the effective force constants of normal modes depend mainly on short range interactions. However, exploiting vibrational spectra depends critically on being able link the modes to a structural model, typically through total energy calculations. Experimental input from structural and spectroscopic techniques and numerical or theoretical input from simulations must therefore be combined in the study of hydrogen-bonds in complex systems.

In this paper, we extend recent work on hydrogen-bonded systems, using a range of analytical and ab initio numerical methods, combined with inelastic neutron scattering, to collagen and its component amino acids in their polymeric, solid state conformations. 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 poly-glycine-II (PG-II) and poly-proline-II (PP-II), super-coiled into right-handed triple-helices. The first synthesis of dried synthetic polypeptides $(Pro–Pro–Gly)_{10}$ was reported by Sakakibara et al. [3]. It contains approximately one water molecule per PPG triplet, equivalent to 7 g water/100 g completely dried $(PPG)_{10}$. Fig. 1 illustrates 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 presence of glycine at every third residue, reducing the steric hindrance and providing an inter-chain hydrogen-bond perpendicular to the helix axis.

Previous studies on hydrogen-bonded molecular crystals showed advances in the use of density functional theory (DFT) methods in the analysis of inelastic neutron scattering (INS) data

* Corresponding author.
E-mail address: johnson@ill.fr (M.R. Johnson).

0301-0104/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemphys.2008.12.005

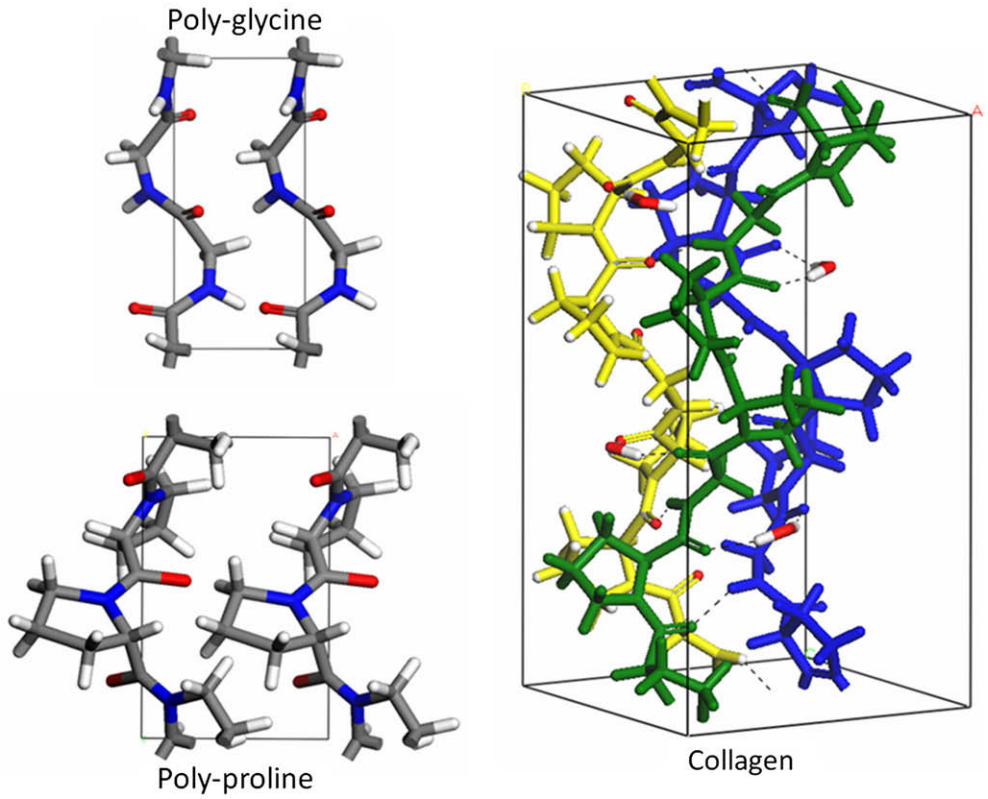

Fig. 1. Helix structures of (top-left) PG-II; (bottom left) PP-II; (right) PPG including some of the water molecules that form hydrogen bond bridges between neighbouring helices. Individual helices are coloured yellow, green and blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[4–9]. In particular, in crystals with short strong hydrogen-bonds, the vibrational analysis based on parameter-free methods revealed the mechanism of the proton transfer as being driven by low frequency phonons, activated by temperature [10,11].

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 [12,13] and it is well known that biological function of a protein is coupled with anharmonic protein dynamics [14–17]. Concerning the study of fibrous proteins, a computational study of collagen based on steered molecular dynamics accessed the elastic properties at the microscopic level [18]. By comparing experimental and simulated INS spectra of PPG_{10}, a model collagen compound, we have carried out a tentative assignment of the lowest frequency vibrations and propose several classes of motion characterizing these elastic properties.

Polymeric systems have been studied with INS and DFT in the case of poly(p-phenylene terephthalamide) (PPTA, known commercially as Kevlar) where the polymeric chains form two-dimensional, hydrogen-bonded sheets in the crystal. Due to the lack of long range order, open questions about the structure remained. Vibrational spectroscopy and quantitative investigation of the amide bands gave important information about the relative orientation of peptide linkages in neighbouring sheets [19].

High frequency vibrations of biopolymers are generally not involved in biophysical processes but they are exploited in the context of the structure-dynamics relation. Thus IR spectra, especially the amide-I band, are commonly used to quantify the presence of beta sheets and helices, dating back to work of Elliott and Ambrose [20], who observed that the frequency of the IR-active Amide I vibration of α polypeptides was about 20 cm^{-1} higher than that of β polypeptides. Further studies on synthetic polypeptides [21] established this observation as a firm empirical rule which can be applied for determining the presence of secondary structures in proteins [22,23]. Topical applications of this method include the study of prion diseases [24].

Finally, in protein science, including the case of collagen, hydration is a very important factor for stability and function [25]. A number of previous studies [26,27] 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 highly dependent on the degree of plasticizing, which is determined by the level of hydration [28]. In this paper, we study the extent to which water vibrations couple to those of collagen.

The purpose of this article is to present results from a range of numerical methods in the study of biologically-relevant, polymeric systems, including crystalline and amorphous water. We have studied poly-glycine in two structural forms (hydrogen-bonded sheets – PGI and helices – PGII), poly-proline (helices – PPII) and hydrated models of synthetic collagen based on the proline–proline–glycine sequence (PPG_{10}). Dry samples of PPG contain one water molecule per PPG triplet and, experimentally, wet samples contain four times this quantity of water. These structures show chemical similarities and topological similarities in that three of them are helical structures. We exploit high frequency spectra to investigate the structure-dynamics relation and the accuracy of solid state, DFT methods in this context. In the case of collagen, the goal is to establish a spectral signature of the triple-helix [29]. By studying dry, homo-polymers and wet, hetero-polymers, a range of computational challenges are encountered, especially with respect to the low frequency dynamics of collagen.