How Were DNA and Protein Molecules Formed During the Process of the Origin of Life?

Xiao-Feng Pang[1,2]
[1]Institute of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, Sichuan, P. R. China
[2]International Centre for Material Physics, Chinese Academy of Sciences, Shenyang, P. R. China

Abstract **-- We propose a new mechanism for formation of DNA and proteins in the process of origin of life based on the discovery of clustered chain of water molecules in water. We suggest that not only was water a suitable environment for appearance of DNA and proteins in primordial times but also that water molecules were their templates of formation. The growth of nucleotide and amino acid molecules along these clustered chains resulted in the formation of the preliminary chain structures of DNA and proteins. This is just one reason why there is not life without water in the world. Our experiments shed new light on the existence of linear and ring hydrogen-bonded chains of water molecules in water. The experiments included infrared absorption spectrum, polarization of molecules, characteristics of first order phase-transition, measurements of heat of vaporization, X-ray diffraction, magnetization of water and self-assembly of linear and helical chains of $(H_2O)_n$. Finally, we propose a mechanism for the specific conformations of DNA and proteins arising from the assembly of nucleotide and amino acid precursors on the templates of these clustered chain structures of water molecules. These results could provide an explanation for the original appearance of the preliminary chain structures of DNA and protein molecules in process of the origin of life. Since the probability forming these preliminary chain structures of DNA and protein molecules on the templates is very small, so, form of life needs very long times.**

Index Terms **-- DNA, Qrotein, Self-assembly, Molecules chain, Protein molecule preliminary structure, Template of water molecular chain, Life birth, Nanomolecule, Water, Hydrogen bond.**

I. BIOLOGICAL BACKGROUND

It is well known that protein and deoxyribonucleic acid (DNA) molecules are the main components required for life; life cannot exist without DNA and protein molecules [1-5]. The molecular structure of protein is well known. Protein molecules are made up of more than twenty different individual building blocks, the amino acids. Each amino acid is composed of an amino group (NH_2), a carboxyl group (COOH), and a side group or radical (R) attached to the α-carbon atom. The nature of the side group distinguishes one amino acid from another. Amino acids polymerise to form the long chains of residues that constitute protein molecules. When two amino acids join together, they release one molecule of water and form a peptide bond. The one-dimensional, linear polypeptide chain is the preliminary structure of the protein molecule. It can fold into a variety of complex three dimensional conformations. Of particular interest are three structural configurations that recur regularly in proteins: the α-helix, the β sheet, and globular conformation [1-5]. In the α-helix the polypeptide chain is tightly coiled about its longitudinal axis. In the β sheet the chain can be visualised as pleated strands of protein. The globular conformation is the most complex since the chain is folded irregularly into a compact, near-spherical shape. The polypeptide chain is most often in the α-helix or β sheet configuration.

In the case of DNA, the molecule consists of four nucleotides, dAMP, dGMP, dCMP and dTMP. Each nucleotide is composed of phosphate, ribose and a base. The bases of DNA are adenine (A), cytosine (C), guanine (G) and thymine (T). Ribonucleic acid (RNA) consists of AMP, GMP, CMP and UMP and its bases are adenine (A), cytosine (C), guanine (G) and uracil (U). These bases link with riboses by means of glycoside bonds. Formation of these bonds involves release of water molecules. Two riboses are combined as a straight chain by the phosphatide bond at the 5' and 3' positions of the neighboring riboses. The preliminary structures of DNA and RNA are one-dimension linear chains. Two chains combine together by hydrogen bonding between complement-ary base-pairs and fold further as duplex and super-helical structures, which form the secondary and tert-iary molecular structures of DNA, respectively [1-5].

From above we know that the preliminary or basic structures of proteins, DNA and RNA are linear

chains. In practice, the molecular structure of the other two major classes of biomacro-molecules, the carbohydrates and lipids, are also linear [1, 2, 4]. This being the case the question arises as to why these preliminary structures are linear. However, alternative structural forms of matter such as face centre cubes, body centre cubes, hexahedrons or other polyhedrons are not formed in living systems. The question then arises of how the linear chains form. At the same time, the above molecular structures of DNA and proteins cannot also achieve their final conformation independently from nucleotide and amino acid molecules since any self-assembly of molecules requires a template, that is to say, DNA and proteins cannot have formed and assembled spontaneously without templates, particularly in prehistory when life was originating. The question of what constituted the original templates for formation of DNA and protein molecules at life's origin remains unanswered. From molecular biology and gene expression rules we now know that proteins are synthesized from amino acid molecules using genetic information supplied by DNA through transcription and translation of RNA in the ribosomes of cells at physiological temperatures in the living systems [1-3, 5].This shows that the proteins can form, only if DNA and RNA exist firstly in the life body. In the meanwhile, the nature of the templates for this process is also unclear. In one sense it is clear that the templates are RNA molecules in the ribosomes. However, the templates for DNA and RNA are not known. Especially in the process of the origin of life there remains uncertainty surrounding the processes by which DNA, RNA and protein molecules formed on earth and what the templates for these original molecules were. Life cannot exist without water, but we do not fully understand the role of water in life's origin. We recently observed the existence of cluster chain structures of water molecules of size 2-70 nm. We contend that water is not only a prerequisite for the biosynthesis of DNA and protein molecules but also for the formation of their preliminary chain structures. We here argue that these cluster chains of water molecules act as templates for the preliminary chain structures of DNA and proteins. Thus we propose further that this is the mechanism of the formation of DNA and protein molecules at life's origin. To substantiate such a claim we have to first present evidence for the existence of linear molecular chain structures in water.

II. THE CHAIN STRUCTURE OF MOLECULES IN WATER

Water is a very familiar material with many unusual properties. For example, water has a large dielectric constant, high viscosity, high surface tension, weak electric conductivity, and low magnetic effect, and it can form laminar flows, turbulent flows, vortices, and various solitary water-waves of bell KdV and nonpropagation and kink [6-8]. These properties are unique and difficult to explain and it is therefore possible that water is composed of more than free molecules of H_2O. Many experiments have demonstrated the complex structure of water molecules and the various unusual properties of water outlined above suggest also that a new structure for the molecule is required. However the nature of such a new structure of water remains unclear.

The infrared spectrum of water provides an insight into its molecular structure. Thus we measured in detail the infrared absorption spectrum of pure water at 35 ℃ using a Nicolet Nexus 670-FT-IR spectrometer. The spectrum is shown in Fig. 1. So-called pure water consists of only water molecules without other impurities and has a pH value of about 7-7.1 which was assessed using some instruments. The experiment was reproduced more than 10 times using pure water collected at different times and in different places and the results were the same as for Fig. 1. We see from this figure that there are three peaks in the region of 300-3800 cm^{-1}. These occur at 300-900 cm^{-1}, 1600-1900 cm^{-1} and 2900-3800 cm^{-1}. The peak at 300-900 cm^{-1} is presumably associated with the liberation of water molecules. The peak at 1600-1900 cm^{-1} is a narrow band arising from bending vibrations of HOH bonds. The peak at 2900-3800cm^{-1} indicates vibration of the OH bond and is a band containing six peaks at 3037, 3165, 3280, 3415, 3540 and 3665 cm^{-1} [9, 10]. This is first time this result has been reported. In order to understand the origin of the six peaks we measured and collected further data on changes to the infrared absorption spectrum of pure water with changing temperature we varied temperature within the range 35-85 ℃ and the results are shown in Fig. 1. The data clearly show that the six peaks remain even with increasing temperature of the water. This indicates that the six peaks reflect the essential structure of water molecules. Drawing on the theory of molecular physics and the ideas of Walrafen et al [11] and Jiang et al [12], the two peaks at 3540 cm^{-1} and 3665 cm^{-1} represent symmetric and antisymmetric stretching vibrations of OH bonds without hydrogen bonds. The two peaks at 3280 cm^{-1} and 3415 cm^{-1} correspond to symmetric and antisymmetric stretching vibrations of OH bonds with hydrogen bonds or hydrogen bonded chains. In the light of the theory of vibration of molecules and their vibrational frequency $v = 1/2\pi \times (K / M)^{1/2}$, where K is the force or elastic constant and M is the effective mass of vibrational

matter, we can confirm that the bands at 3280 cm^{-1} and 3415 cm^{-1} are the result of a red-shift frequency of 3540 cm^{-1} and 3665 cm^{-1}, respectively, due to the increase of vibrational mass when the hydrogen-bonded chains occur. In general, the greater the number of water molecules involved, the larger the red-shift frequency of the peaks although the shift is not proportional. Our own and other [9, 10] experiments have also found that the two peaks at 3540 cm^{-1} and 3665 cm^{-1} almost disappear in ice crystals and that the two peaks at 3280 cm^{-1} and 3415 cm^{-1} also disappear in vaporised water. We can therefore confirm that the above assignment is correct. Thus we believe that there are not only a large number of free molecules but also that many hydrogen bonded chains of molecules occur in water.

The above experiment using infrared spectroscopy suggests the existence of linear hydrogen bonded chains in water. However, the issue of how free water molecules form these hydrogen bonded chains is unclear. As is well known, when two hydrogen atoms and one oxygen atom with high-electronegativity combine to form a water molecule, the two hydrogen atoms and their charges are not symmetrically distributed relative to the oxygen atom; there is always an angle between the two OH bonds which is about 105 ℃ (it alters to 109 ℃ in water). This shows that each water molecule is polarized and has an electric dipole moment [4, 5, 13]. This state can easily be demonstrated experimentally by the movement of a stream of water towards a statically charged plastic rod. Experiment and calculation show that each water molecule has a large electrical dipole-moment of 1.84 Debyes (the dipole moment of an OH group is only 1.5 Debyes) [13]. Thus there are certainly a large number of hydrogen-bonded chains in water molecules. For instance, dimers $(H_2O)_2$ and other polymeric complexes are associated with hydrogen bonds of length of 0.276 nm and bond energy of 0.31-0.34 eV, where the covalent bond is 0.11 nm, and the weak long-bond is 0.67 nm. Therefore, each hydrogen-bonded chain containing several hundred molecules has a size of nanometer scale [9, 10, 14, 15].

The new structure of molecules with hydrogen bonded chains cannot be destroyed by thermal perturbation and spatial fluctuation in water in the range of 0-100 ℃ because the hydrogen-bonded energy is more than an order magnitude larger than the thermal energy of the molecules (5 KT/2= 0.041-0.05 eV from 0-100 ℃), thus it can suppress the destructive effects of these factors on the hydrogen bonds [5]. This is verified by the experimental result shown in Fig. 1. From this figure we see clearly that the six peaks in the infrared absorption spectra are maintained up to the higher temperature of 85 ℃, but the number of associated water molecules in the hydrogen- bonded system and its density possibly decrease with increasing water temperature. Evidently, this is due to thermal disruption of the motion of the molecules. Rough estimates suggest that the hydrogen bonded chains are composed of 240 molecules at room temperature, 150 at 37 ℃, 120 at 45 ℃ and 40 at 98 ℃ [5, 9, 10, 13] with a size of around 10-70 nm, respectively. We can therefore confirm that there are always many hydrogen bonded chains in water [8, 9], whether at high temperature or at low temperature, although the number of molecules in these chains vary with temperature.

On the other hand, at the vaporization point, the liquid water becomes fully vapour at 100 ℃ by absorbing vaporized heat energy of 9.715 kcal/mol = 0.42 eV, which is larger than the energy of the hydrogen bond (0.31-0.34 eV), but an order of magnitude lower than the energy of covalent bonds [5, 9, 10]. Therefore, the energy absorbed can destroy all hydrogen bonds, but cannot destroy the covalent bonds of water molecules. Thus these molecules in hydrogen-bonded chains in water can become free molecules in the vapour phase. The residual energy {(0.42-0.31) eV = 0.11 eV or (0.42-0.34) eV = 0.08 eV} is used as kinetic energy to disrupt the motion of vapour molecules. In such a case the volume of water increases about 1660 times from V_{liq} = 1.04346 cm^3/g to V_{vap} = 1673 cm^3/g for 1 g water at 1 atm and 100 ^oC, where dp/dT = 0.0356 atm/℃ and p is pressure of water [4, 5, 9, 10]. Thus we can only observe the two peaks at 3540 cm^{-1} and 3665 cm^{-1}, associated with the free molecules in the infrared range in the region of 3000-3700 cm^{-1} in water vapour. These results have been verified by many experiments [9, 10, 14, 15]. These experimental results also lead us to believe that there is always a large number of hydrogen bonded chains in water from 0 to 100 ℃.

From the above experimental results with pure water we can confirm that there are not only a great number of free molecules but also many linear hydrogen-bonded chains in water. The linear and helical hydrogen-bonded chains of water molecules containing $(H_2O)_{10}$, $(H_2O)_{12}$ and $(H_2O)_{15}$ have already been self-assembled by artificial experiments to produce in organic super-molecular complexes [16-20]. These results again confirm and verify experimentally the linear and helical hydrogen-bonded chains of molecules in water. We therefore have sufficient reason to believe that the linear hydrogen-bonded chains of molecules are an essential and intrinsic feature of water.

III. FORMS OF DNA AND PROTEIN MOLECULES WITH CHAINS IN THE PROCESS OF THE ORIGIN OF LIFE

It is well known that water is an important environmental constituent that was necessary for the appearance of life on earth [3-5]. There appear to have been only water, amino acid molecules or nucleotide molecules on earth at the time of life's origin. Water provides a suitable environment for the development of life including the formation of bio-tissues, DNA and protein molecules. DNA and protein molecules may have proliferated gradually from water but some kind of template must have existed when these nucleotide and amino acid molecules formed and self-assembled as the preliminary structures of the one-dimension chains of DNA and proteins. We suggest that molecular chains within water facilitated the arrangement of amino acid molecules or nucleotide molecules along the backbone of water molecule chains. According to our theory, these chains of water molecules were therefore just the templates for formation of DNA and protein molecules from nucleotides or amino acids, respectively, in the process of the origin of life. The precise mechanism underlying this phenomenon deserves further study.

As we known, amino acids and nucleotide molecules have an electrical dipole moment of 4-6 Debyes and 5-6 Debyes [5, 15], respectively. Each water molecule and O-H bond also have a dipole moment [5, 15], respectively. When they are together in a system, amino acid and nucleotide molecules move toward water molecular chains due to the dipole-dipole attraction interactions among them, combining together via hydrogen bonds between the hydrogen atoms in the O-H bonds of water molecules and the oxygen atoms of amides (O=C) in these amino acids or the oxygen atoms of ribose rings in nucleotide molecules. Thus these amino acids and ribose rings in the nucleotide molecules could then be fixed onto the templates of water molecular chains through hydrogen bonds as shown in Fig. 2-5. Once a number of amino acids had been fixed onto the templates of water molecular chains (Fig. 2), these amino acids could then form the preliminary structures for the one-dimensional chains of protein molecules by peptide bonds between neighbouring amino acids, where one water molecule is released as shown in Fig. 3. Subsequently, they could fold to form the α-helix, β sheet and globular conformations (or secondary, tertiary and quaternary structures) of protein molecules with the support of various hydrogen-bonds which would be generated under appropriate conditions of water environment and temperature. This is our proposed mechanism for the formation of protein molecules from amino acid molecules in the process of life's origin. In the same way DNA could also have been formed. Thus, after many ribose rings in the nucleotide molecules had been fixed onto the templates of water molecular chains through hydrogen bonds, these nucleotide molecules could subsequently form the preliminary structures of the one-dimensional chains of DNA by phosphate bonds between the 5' position of ribose rings and the 3' position of neighbouring ribose rings with the release of two water molecules. At the same time, one base could also link to the 1' position of the ribose ring with the release of one water molecule (Fig. 4). In practice, the preliminary structures of DNA could also be formed in another way: the phosphate molecules could firstly be fixed onto the templates of water molecular chains through hydrogen bonds between the O-H bonds of water molecules and oxygen atoms of 0═P bonds in the phosphate groups. Subsequently these ribose rings could be linked via these phosphate groups by phosphate bonds to form the preliminary chain structures of DNA (Fig. 5). Finally, both one-dimensional structures could form a double-chain of DNA and fold further to form the higher structures with their helical or super-helical shapes. The above mechanism and process for the formation of preliminary chain structures of DNA and protein molecules may have been the processes occurring at the origin of life and the route to formation of DNA and proteins from their nucleotide and amino acid precursors, respectively. From consideration of these processes we suggest that the water molecular chains may have played a key role in the origin of life. The chain structures of water molecules are simply the templates for the preliminary structures of the one-dimensional chains of DNA and protein molecules.

We also know that the processes of formation of preliminary chain structures of DNA and protein molecules are completely spontaneous. As mentioned above, in the early processes of life's origin there were only water molecules, amino acids and nucleotide molecules. Under such conditions, the above mechanism and process are the only way in which DNA and protein molecules could have been formed. Thus we can say that no preliminary chain structures of DNA and protein could have formed without the linear water molecules in the process of the origin of life. We also hypothesise that there would be no DNA and protein molecules or life on earth without the prior presence of water on earth. This is just one reason why there is no life without water in the world because DNA and protein

molecules are main and basic components of life. This is a novel idea which has not previously been proposed. The mechanism of the formation of DNA and protein molecules does contradict the theory of molecular biology, in which the positions of amino acids on the proteins are controlled by genetic information of DNA in life systems [1-4]. However, this theory of gene expression of DNA has not solved the problem of how these amino acids are only arranged as chains and not in other forms. Since the probability forming these preliminary chain structures of DNA and protein molecules on the templates is very small, thus we can also explain why form of life needs very long times by this theory.

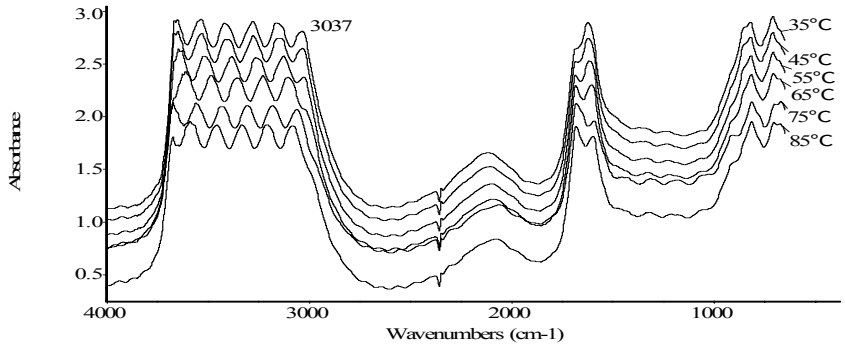

Fig.1. The infrared absorption spectra of pure water and their frequency shifts at different temperatures T = 35 ℃, 45 ℃, 55 ℃, 65 ℃, 75 ℃, 85 ℃.

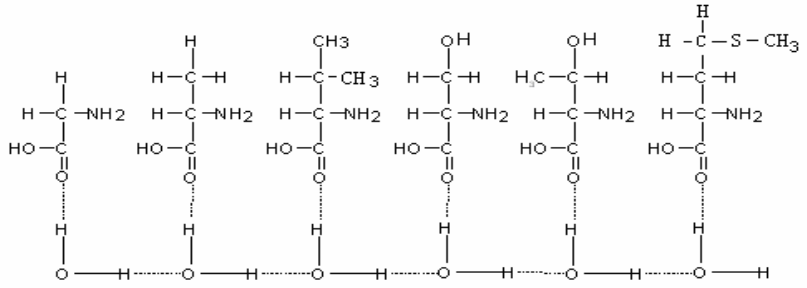

Fig.2. The situations of distribution of amino acid molecules of G, A, V, S, T, and M fixed on the templates of water molecular chains.

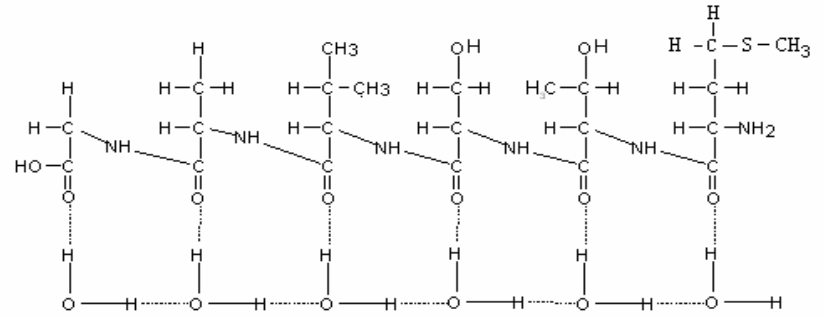

Fig.3. A form of single proteinic chain built by the amino acid molecules of G, A, V, S, T, and M on the basis of templates of water molecular chains.

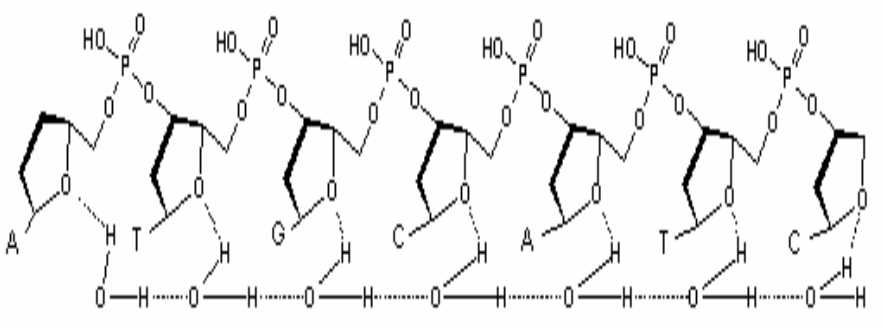

Fig.4. A form of single DNA chain built on the basis of templates of water molecular chains, where A, T, G, and C represent four bases in DNA, respectively.

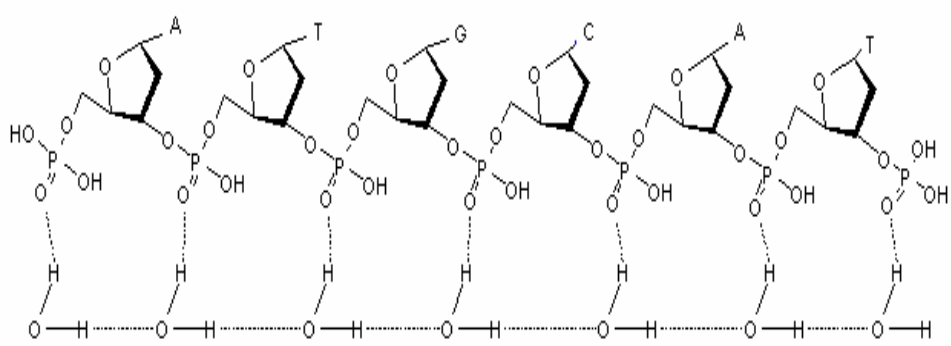

Fig.5. Another form of single DNA chain built on the basis of templates of water molecular chains, where A, T, G, and C represent four bases in DNA, respectively.

In the early processes of life's origin there is, as yet, no known sequence for the appearance time of DNA or protein molecules. In fact, we see clearly from the above processes that the forms of DNA and protein molecules are independent and not mutually dependent as previously thought. In other words the formation of DNA and protein molecules may have occurred independently and simultaneously during the process of life's origin. Therefore we can't conclude that DNA arose before proteins, or vice versa. Thus we can solve the long-time controversial problem of the origin sequence of DNA and protein molecules in life science. Finally, it follows that not only is there no life without water but also that protein molecules and DNA may have first appeared simultaneously, taking their origins from water.

ACKNOWLEDGEMENT

The author would like to acknowledge the National "973" project of China for financial support (2007CB936103).

REFERENCES

[1] B. Alberts, D. Bray, J. Lewis, et al, Molecular biology of the cell, 3rd ed, New York, Garland Publishing, Inc., 1994.

[2] J. H. Lodish, D. B. altimore, A. Berk, et al, Molecular cell biology, 3rd ed., New York, Sceintific American books, 1995.

[3] X. F. Pang, Phys. Rev. E, vol. 62, pp. 6898, 2000.

[4] (a). X. F. Pang, Y. P. Feng, H. W. Zhang, and Y. H. Luo, J. Phys. Condensed matter, vol. 18, pp. 613, 2006; (b). X. F. Pang, Biological electromagnetics, Chinese Defect Industry Press, Beijing, 2007.

[5] X. F. Pang, Biophysics, Press of Univ. of Electronic Sci. Tech. of China, Chengdu, 2007.

[6] (a). R. Z. Sagdeev, Nonlinear and turbulent processes in physics, Harwood Press, 1984; (b). S. D. Liu and S. K. Liu, Solitary wave and turbulence, Shanghai Sci. Tech. Edu. Publishing House, Shanghai, 1994.

[7] (a). J. R. Wu, R. Keolian, and I. Rudnick, Phys. Rev. Lett., vol. 52, pp. 1421, 1984; (b). B. Denardo, W. Wright, and S. Putterman, Phys. Rev. Lett., vol. 64, pp. 1518, 1990.

[8] X. F. Pang, Soliton physics, Sichuan Sci. Techno. Press, Chengdu, pp. 1-253, 2003.

[9] (a). X. F. Pang, European Phys. J. B, vol. 49, no. 5, 2005; (b). X. F. Pang, J. Chin. Atom. Mol. Phys. (supp.), vol. 23, pp. 1, 2006.

[10] (a). X. F. Pang, and B. Deng, Physica B, vol. 403, pp.3365, 2008; (b). X. F. Pang, and B. Deng, Sciences in China G, vol. 51, pp. 1357, 2008; (c). X. F. Pang, and B. Deng, Chin. J. Atom. Mol. Phys., vol. 24, pp. 281, 2007; (d). B. Deng, and X. F. Pang, Chin. Science Bulletin, vol. 52, pp. 999, 2007; (e). X. F. Pang, and Y. P. Feng, Quantum mechanics in nonlinear systems, World Science Publishing Co. New Jersey, pp. 577-581, 2005.

[11] (a). G. E. Walrafen, J. Chem. Phys., vol. 36, pp. 1035, 1962; (b). G. E. Walrafen, J. Chem. Phys., vol. 40, pp. 3249, 1964; (c). G. E. Walrafen, J. Chem. Phys., vol. 44, pp. 1546, 1966; (d). G. E. Walrafen, J. Chem. Phys., vol. 48, pp. 244, 1968; (e). G. E. Walrafen, J. Chem. Phys., vol. 50, pp. 560, 1969; (f). G. E. Walrafen, J. Chem. Phys., vol. 52, pp. 4276, 1970; (g). G. E. Walrafen, M. S. Hokmabadi, and W. H. Yang, J. Chem. Phys., vol. 85, pp. 6964, 1986; (h). G. E. Walrafen, M. R. Fisher, M. S. Hokmabadi, and W. H. Yang, J. Chem. Phys., vol. 85, pp. 6970, 1986.

[12] Y. J. Jiang, Q. J. Jia, P. C. Zhang, and L. Xu, J. Light Scattering, vol. 4, pp. 102, 1992.

[13] A. S. Davydov, Biology and quantum mechanics, New York, Pergamon, 1982.

[14] (a). X. F. Pang, and H. J. W. Muller-Kirsten, J. Phys. Condensed Matter, vol.12, pp. 885, 2000; (b). X. F. Pang, and Y. P. Feng, Chem. Phys. Lett., vol. 373, pp. 392, 2003; (c). X. F. Pang, and A. F. Jalbout, Phys. Lett. A, vol. 330, pp. 245, 2004.

[15] X. F. Pang, and Y. P. Feng, Quantum mechanics in nonlinear systems, World Scientific Publishing Co. Singapore, pp. 557-586, 2005.

[16] L. J. Barbour, G. W. Orr, and J. L. Atwood, Nature, vol. 393, pp. 671, 1998.

[17] M. Yoshizaawa, T. Kusukawa, M. Kawano, et al, JACS, vol. 127, pp. 2798, 2005.

[18] S. K. Ghosh, and P. K. Bharadwaj, Angewandte Chem., vol. 43, pp. 3577, 2004.

[19] B. Sreenivaasulu, and J. Vittal, Angewandte Chem., vol. 43, pp. 5769, 2004.

[20] (a). S. Y. Yu, H. Hui, H. B. Liu, et al, Angewandte Chem., vol. 42, pp. 686, 2003; (b). S. Y. Yu, Z. X. Zhang, C. C. Eddie, et al, JACS, vol. 127, pp. 17994, 2005.