

## THEORETICAL MODELLING OF DNA BASE SUPER-STRUCTURES ON THE GOLD (111) SURFACE

R. E. A. Kelly, L. N. Kantorovich

Physics Department, King's College London, London, WC2R 2LS, U.K.

ross.kelly@kcl.ac.uk

The DNA double helix, one of the most amazing and important self-assembled structures found in nature is dependent on the interaction between the DNA base molecules: Adenine (A) and Thymine (T); and Cytosine (C) and Guanine (G). Recently there have been a number of Scanning Tunneling Microscope (STM) experiments on the gold (111) surface in which the DNA base molecules (Adenine (A), Cytosine (C), Guanine (G) and Thymine (T)) were investigated [1-3]. The STM images show very distinct super-structures of these molecules on the surface. Whereas Adenine [1] and Guanine [1,2] show amazing self-assembled periodic monolayers in the form of flowerbeds, Thymine [3] and Cytosine [1] form long filaments, which are very mobile at room temperature.

The self-assembly of organic molecules is an important research field due to potential uses of self-assembly within the realms of nano-technology [4], with a strong possibility of 2D 'supramolecular architectures' being formed on solid surfaces. From a biological view-point, it has been suggested [5] that self-assembly of small organic bases at the solid-liquid interface may provide an explanation for possible mechanisms of pre-biotic evolution when sophisticated organic molecules were formed.

Unfortunately, it is not possible at present to resolve the atomistic structure of the monolayers experimentally, so that only theory can do this. A simple analysis that we present which is based on *ab initio* calculations in conjunction with certain assumptions concerning the way the molecules are connected in the super-structures, reveals that in some cases there are many different arrangements of the molecules possible, which can explain the observed STM images.

Using calculations based on the *ab initio* Density Functional Theory (DFT), we for the first time report *all* possible planar DNA base homo-pairs. Our analysis is based on a systematic identification of all *binding sites*, or simply *sites*, where one molecule can bind to the other by a double or triple Hydrogen bond. We find many possibilities for each molecule, ranging from 10 for Thymine to 17 for Guanine, most of which have never been considered in the previous literature. Even some important high stability homo-pairings have been previously neglected, which we consider as important in forming some of our predicted super-structures. All possible pairings between different monomers have also been identified and studied.

We also found that stability of a Hydrogen bond can be characterised by the difference charge density that shows well-developed regions of alternating excess and depletion of the electron charge similar to a "kebab" structure. Fig. 1 shows examples of the difference charge density for a strongly (a) and weakly (b) bound A-A pair.

The presented detailed information on all planar DNA base homo-pairs is utilised in predicting all possible super-structures based on the DNA bases. In particular, we have studied all possible Adenine hexagonal networks (super-structures with a three-fold coordination of each monomer) and Cytosine filaments (1D chains with two molecules in the unit cell). Two examples of super-structures are shown in Fig. 2. Note that in our method no

DFT calculations are initially required for the predicted super-structures, as their geometry and stabilisation energy are obtained exclusively from the DFT calculations of the corresponding dimers.

Depending on the substrate, different super-structures have been observed experimentally, and if one can assume that there is small interaction with the substrate, all these structures should be predicted in our study. Using this information we can suggest how many suitable candidates there are for the various super-structures. The cases of the Adenine and Cytosine super-structures are considered in more detail.

We find that thousands of energetically favourable Adenine monolayers are possible, just a few have been studied previously. Full relaxation DFT calculations applied to a finite set of randomly selected gas-phase monolayers showed a good correlation with the predicted structures and binding energies from our method (see Fig. 2). Similarly, we find many tens of Cytosine filaments are possible with a possibility for some of them to form rings and filament connections that have also been seen in some of the experimental images [1].

Careful analysis of the stabilisation energies for the super-structures indicates the importance of the cooperative effects in the binding of the molecules that represent a hydrogen bond 1D or 2D networks (cf. [2]).

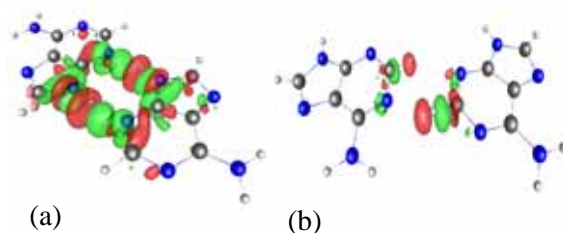
We also question in our work the suitability of semi-empirical and semi-classical methods in studying DNA bases super-structures.

By considering the DNA bases, we suggest a vast number of possibilities for their super-structures. One can expect that a similar situation may also happen in the cases of other molecules forming several hydrogen bonds at the same time in the network. Therefore, one should be careful in suggesting a particular model for the observed super-structures (e.g. from STM experiments) as many more structures may well be possible. To pin-point to the particular structure, all these possibilities have to be analysed using the method similar to that suggested in our work and should then be followed by more detailed *ab initio* calculations.

### References:

- [1] R. Otero et al., Phantoms Newsletter, 17: 5, 2004.
- [2] R. Otero et al., Angew. Int. Ed. 44: 2270-2275, 2005.
- [3] R. Otero et al., Private communication.
- [4] S. De Feyter, C. De Schryver, Chem. Soc. Rev. 32:139-150, 2003.
- [5] S. J. Sowerby, W. M. Keckl, Orig. Life and Evol. Biosp. 28:283-310, 1998.

**Fig. 1**



**Fig. 2**

