Zipper-like structures, alternative conformations of DNA in guanine rich regions and importance of cations for their stability

Miguel Angel Méndez[1,2]*, **Andrea Montero**[1,2], **Ana Samaniego**[1,2], **Andrea Sosa**[1,2]
Dennisse Vallejo[1,2], **Silvia Velasteguí**[1,2], **Francisco Yanqui-Rivera**[1,2]

[1] *Universidad San Francisco de Quito, Colegio de Ciencias de la Salud,*
Vía Interoceánica Km 2 1/2 y Av. Florencia, Quito, Ecuador.
[2] *Grupo Ecuatoriano para el Estudio Experimental y Teórico de Nanosistemas –GETNano–*
Diego de Robles s/n y Vía Interoceánica, Quito, Ecuador
Autor principal/Corresponding author e-mail: mmendez@usfq.edu.ec

Abstract

The manufacture of nanostructures from bio-molecules such as DNA strands is an ongoing exploration frontier by a multitude of potential applications in biological complex systems. The analyzed sequences are found in biological systems as origins of replication of viruses and bacteria. In this paper we report the detailed characterization at the atomic level of Guanine-rich DNA that forms zipper-like structures by molecular mechanics modeling. Models were constructed and these structures minimized, allowing analysis via molecular mechanics to understand the factors that determines the most stable structure. It was found that the presence of positively charged ions near the Gs-rich region of the studied sequences, is critical to the stability of these DNA structures. In summary, the results allow a better understanding of this system at the molecular level allowing the development of more efficient procedures for the control of the manufacture of zipper-like DNA nanostructures and finding their applications in biological systems.

Keywords. guanine rich DNA sequences, zipper-like structures, molecular mechanics, nanotechnology

Resumen

La fabricación de nanoestructuras a partir de biomoléculas como fibras de ADN es una frontera en continua exploración gracias a una multitud de aplicaciones potenciales en sistemas biológicos complejos. Las secuencias analizadas se encuentran en sistemas biológicos como en orígenes de replicación de virus y bacterias. En este artículo se reporta la caracterización detallada a nivel atómico de ADN rico en Guaninas que forman estructuras en cremallera. Para esto se utilizaron métodos de modelación con mecánica molecular. Se construyeron modelos para las estructuras y se minimizaron, lo cual permitió su análisis vía mecánica molecular para entender los factores que determinan la estructura más estable. Se encontró que la presencia de iones con carga positiva cercanos a la región rica en Gs de las secuencias estudiadas es de importancia fundamental en la estabilidad de estos ensambles de ADN. En resumen, los resultados permiten mejorar la comprensión de este sistema a escala molecular permitiendo desarrollar procedimientos más eficientes para el control de la fabricación de nanoestructuras en ADN en cremallera y encontrar sus aplicaciones en sistemas biológicos.

Palabras Clave. secuencias de ADN ricas en guanina, estructuras tipo cierre, mecánica molecular, nanotecnología

Introduction

Nanotechnology is exploring a new field directed to the manufacture of objects, devices and materials starting from bio-molecules with potential applications in complex biological systems [1] DNA is one of the most versatile molecules that are being explored for applications in this area. These processes, implies the usage of DNA in scaffold materials, nano-electronic components and polymers for the formation of biological nanostructures and nano-fibers. The nanofibers have become a topic of great interest in recent years due to its enormous potential in many areas, especially in biomedical applications [2]. DNA has emerged as a versatile ma-

terial to build artificial molecular structures with excellent intrinsic properties which include programmability, self-organization, molecular recognition, and molecular scale structure fabrication. Therefore, these properties make DNA attractive as a nano-scale construction material [3]. The DNA has the capacity of self-assembly making it into a natural candidate to form nanostructures, leading and directing their self-assembly. This is the principle whereby the DNA is a bio-molecule especially important in the field of nanotechnology [3]. In this paper, it will be studied a non-conventional DNA: DNA in a zipper-like folding. This conformation of DNA has been found in naturally occurring sequences and it has been observed that it has a very stable structure [4, 5]. Some bacteriophages present in their origin of replication a hairpin loop with a rich purine sequence, same as repetitive sequences found in centromeric DNA that can form a structure zipper-like type *in vitro* [4]. The stability of these structures is directly related to the formation of hairpins or unusual duplexes [4]. These have been repeatedly studied through nuclear magnetic resonance (NMR) indicating that the sequence of origin of replication of bacteriophage usually forms hairpins or loops, while the same sequence in centromere results in duplex formation and stem loops [4].

Methods

Through molecular dynamics modeling it is intended to establish whether a sequence of DNA rich in guanine would self-assemble in a DNA with a zipper-like structure. A model structure for these sequences is minimized iteratively utilizing the program NAMD (Theoretical and Computational Biophysics Group – University of Illinois at Urbana-Champaign) in order to optimize the geometry of the structure [6]. In this process once you get the three dimensional structure of the modeled sequence, the program proceeds to relax the tensions within the structure so that the atomic positions are reasonable. This requires several steps of energy minimization [6]. The series of experiments described here seeks to make an atomic-level characterization of this model system analyzing various structures of duplexes with a central region of base-pair mismatches, rich in guanine to obtain principles that can be applied to understand under what conditions these molecules can form alternative structures such as zipper-like on biological systems [7].

Preparation of the Model

The zipper-like type structure was created in the program 3DNA, where initially we copy the coordinates of a normal guanine then the bases were stacked manually in the form of a zipper [8]. For this we had to change a number of properties in each guanine, such as helical twist, distance between bases, etc. We rely on the structure parameters for 2FZA crystal structure available in the Protein Data Bank [9]. This was done so that the Gs would not overlap in the same coordinates and

they would be distributed in a zipper-like structure. After preparing a structure with 8 Gs, it was loaded twice in Visual Molecular Dynamics (VMD) software package (Theoretical and Computational Biophysics Group – University of Illinois at Urbana-Champaign) and through translations and rotations of both molecules a 16 Gs stacked structure (zipper-like DNA) was achieved. This structure was saved in PDB file format that stores the structural position of all atoms, and constituted the initial zipper-like DNA model.

Two models were created for two similar sequences, SQ1 and SQ2 that differ in the sequence of the B-DNA duplex ends. Both models were prepared using the 3DNA program parameters 2.0 for Windows [8]. Moreover, visualization and analysis was performed using VMD (Visual Molecular Dynamics) for Windows version 1.9 (University of Illinois-Urbana Champagne).

SQ1 and SQ2 both were introduced in a water box with boundary dimensions that are at least 12 Å greater than the position from any atom of the DNA molecule. The size of the water box ensures that there is no interaction between the molecules of DNA and that during the simulation the molecule does not exit the box of the solvent. Subsequently, we proceeded to neutralize the charge of the DNA molecule with sodium ions following the protocol of auto-ionization, an application of VMD that sets water molecules randomly distributed with a minimum solutes distance of 5 Å and between ions also of 5 Å. It was also changed the position of various ions by moving them from the solvent and getting them closer to the guanine-rich zone since the structures of zipper-like DNA, from which there is experimental structural information, include cations close to the bases in the zippe-like zone [9]. With this change it is expected to see the effects of cations in the structures analyzed in this study and to see if these cations help stabilizing these present structures allowing them to find the most stable geometries.

Minimization

The minimization of the molecules was conducted in several stages. The first step included only the solvent (water and sodium ions of the solvent) such that the DNA atoms were left fix together with the ions in close proximity to the Gs-rich region; this system was minimized by 3000 steps. In the second phase the solvent was minimized including the ions found in the region rich in Gs by 3000 steps. Finally, the whole system was minimized by 17000 steps. This process was carried out using NAMD Scalable Molecular Dynamics 2.8 (University of Illinois-Urbana Champagne) for Linux installed on a cluster of 12 processors (Linux Rocks distribution 5.4.3 (viper)). For the geometry optimization the parameters and topology such as found in top_all27_prot_lipid_na.inp and par_all27_prot_lipid_na.inp files were used. A 10 Å cutoff was applied to the non-bonded interactions, the non-bonded pair list was updated every 10 steps. The particle mesh Ewald method (PME)
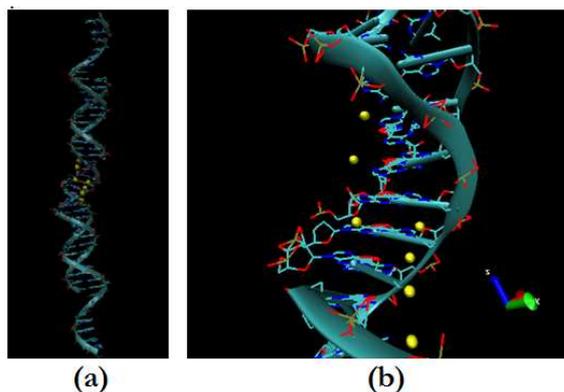
Figure 1: (a) Visualization in New Cartoon in VMD of sequence SQ2. (b) Zoom of the Gs rich region with ions Na$^+$ added for stabilization of the structure.

was used thorough the simulations. To speed up the fast Fourier transform in the calculation of the reciprocal sum, the size of the PME charge grid was chosen to be a product of powers of 2, 3 and 5 [10].

Results

For the construction of models of SQ1 and SQ2 structures we rely on modified sequences from those originally used by Venczel in the formation of DNA synapsable [11]. In our models the sodium ions were placed manually with a VMD program implementation, they are distinguished as golden spheres in the region rich in Gs (Figure 1).

During the minimization of the SQ2 sequence it was observed that the ions moved from their positions. There were changes in the arrangement of the guanines as well (at the area where the ions are added) in comparison with the minimization of the same structure without ions in the region analyzed (Figure 2 and 3). In Table 1 the distances at which the ions are placed on the SQ1 sequence at the end of the minimization are reported. Namely, the presence or absence of ions influences significatively the final optimized geometry of the analyzed sequences. Geometry optimizations for the sequences SQ1 and SQ2 were performed by a total of 25000 and 29000 steps respectively, which assured the structures reach a minimum in energy. To check if the structure is not changing anymore certain parameters were used as the total energy of the system (Figure 4), the RMSD (root media square displacement) (Figure 5). The minimization of unconventional DNA structures zipper-like type came to converge; we see that the energy of the optimized geometry of the two molecules reached a constant energy way before the final steps of minimization. (Figure 4)

We can see in Figure 5 and Figure 6 that the Gs with the most significant changes are the ones in the guanine rich region (the zipper-like structure in these particular models). In these zipper-like structures the Gs are not normally connected (by hydrogen bonds) with its complementary base, but are overlapped one upon the other.
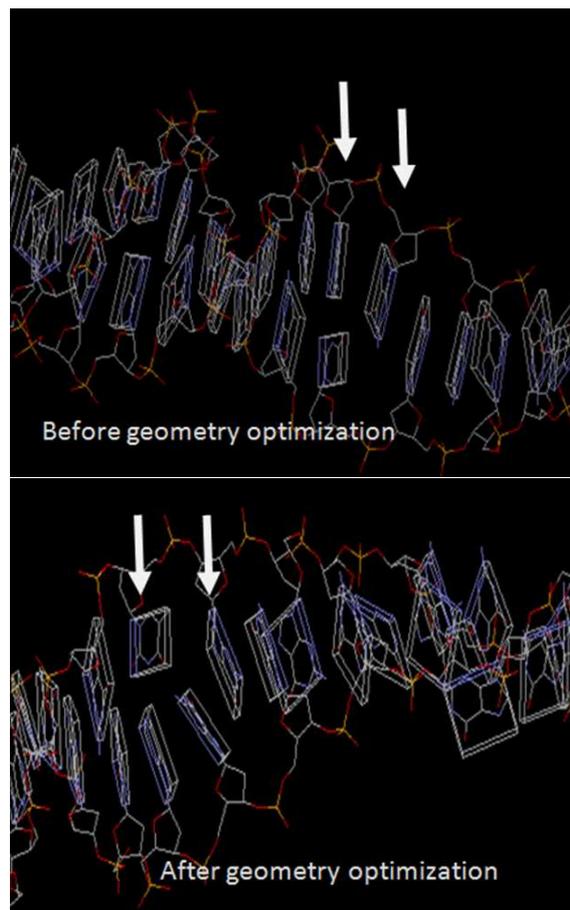

Figure 2: Block diagram of the sequence SQ1 (with ions) before and after minimization. The arrows indicate one of the two regions of transition between double-stranded region and the zipper-like zone.
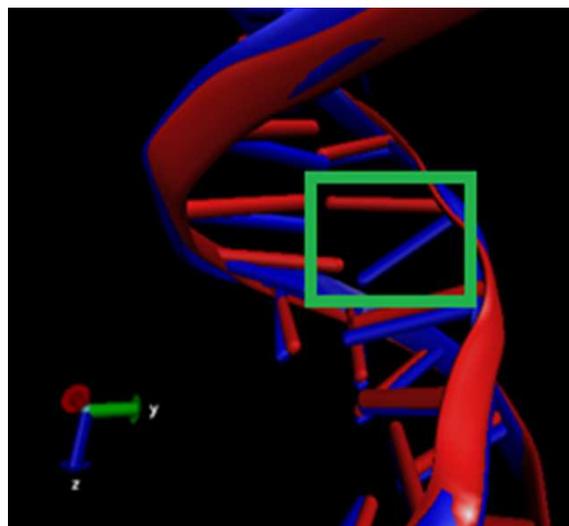

Figure 3: Comparison of Gs-rich region in the sequence SQ2 with Na$^+$ (red) and without Na$^+$ (blue). The green frame shows clearly changes in the position of the base.

Discussion

According to *in silico* experiments and laboratory observations, it is observed that the structural form, topology and therefore function of a biological system depends
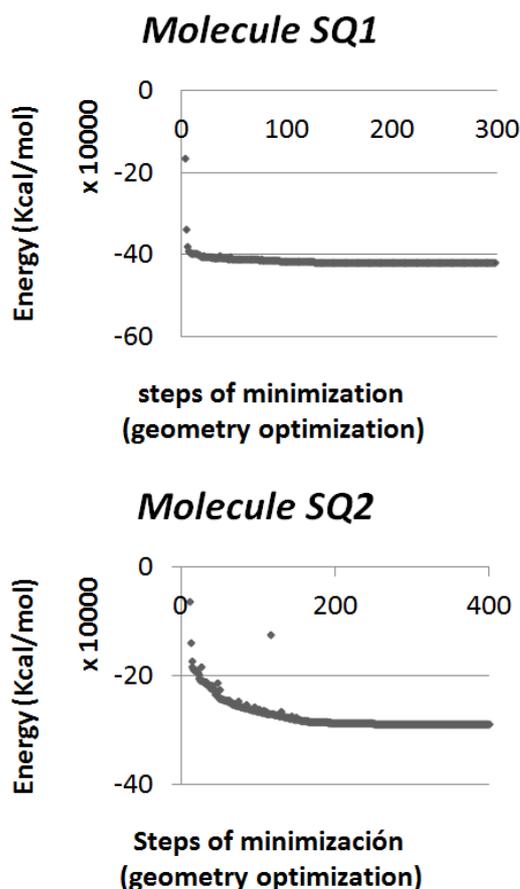
Molecule SQ1

Molecule SQ2

Figure 4: Plots of the total energy vs steps of minimization for the structures after minimization (geometry optimization). (a) The SQ1 sequence. (b) SQ2 sequence. In both cases a minimum in energy has been reached.

greatly on the environmental conditions were it is found [12]. The DNA molecule is rich in negative charges due to phosphate backbone, therefore it is of great importance the addition of cations to stabilize the system and reduce the free energy of hydration [13]. For example, the presence or absence of ions in the surrounding of a DNA molecule can cause it to be stacked (or folded) in a different way, such is the case for G-quadruplex DNA configurations [14]. Thus, one can propose the possibility of controlling the fabrication of different DNA structures from the same or similar sequences but under different solution composition (i.e. different ionic strength) that will influence how DNA stacks, compact, or fold at the nano-scale. Therefore, changing the structure we can change the properties of molecules, for example, their electronic conductivity, with possible applications in the electronics field [5, 15, 16, 17].

There are specific interactions between ions and DNA which are essential in many biological processes [18]. In this work it can appreciated the functionality of the ions in stabilizing the three dimensional structure of DNA in both molecules. This is shown in Figure 1, in which the sodium ions are visible in the zipper-like region; in

			Inicial (Å)	Final (Å)
Gua 19	O6	Na 1	2.28	2.36
Gua 64	O6	Na 1	6.2	6.78
Gua 20	O6	Na 2	2.29	2.25
Gua 63	O6	Na 2	4.06	3.9
Gua 21	O6	Na 3	2.29	2.25
Gua 62	O6	Na 3	4.12	5.08
Gua 22	O6	Na 4	7.03	9.01
Gua 61	O6	Na 5	4.65	7.86
Gua 23	O6	Na 5	2.32	2.29
Gua 60	O6	Na 6	6.06	7.2
Gua 24	O6	Na 6	2.4	2.26
Gua 59	O6	Na 6	5.28	5.29
Gua 58	O6	Na 7	7	5.74
Gua 25	O6	Na 7	5.27	5.89
Gua 57	O6	Na 8	6.2	6.01

Table 1: Distances measured from O6 from a guanine to the closest sodium cation.

Figures 2 and 3, we observe how during the minimization of the system, the sodium ions move into the space directly interacting with the DNA until atoms are accommodated in a geometry where the molecule and ions are stable (geometry optimized). (See Figures 2 and 3 and Table 1)

Because these structures (unusual DNA structures) are more common in the presence of alkali metal cation, sodium was used as the stabilizer cation [19]. There is an important biological meaning in the addition of ions as they cause conformational changes to the DNA molecules. In addition these experiments *in silico* comply with the premise that the DNA can have different structures depending on the solution in whish it is palced [20, 21]. The effect of the cations in the structure of the zipper-like SQ2 (Figures 2 and 3), is the neutralization of a significant negative charge in the central area of the structure. In Figure 5 and 6 the structural changes in the zipper-like area and in the duplex region are compared. The amount of changes in the structures shows that the effect of adding ions is greater in the structure of the zipper-like region that in the duplex B-DNA region. By having base to base stacking interactions in the structure of the zipper-like DNA, charged atoms (such as sodium cations) are needed to neutralize the molecule, thus obtaining a system with lower free energy after minimization. However, it is observed that in a conventional duplex of DNA (Figure 6), the structure is stabilized quickly without the presence of cations placed a priori in proximity to the bases; but it does so only by the presence of ions in the aqueous medium (water box) of the molecule positioned using the VMD auto-ionization tool.

It is noteworthy that the final conformational structure of the molecule was able to be achieved by geometry optimization only when the cations were placed in a relatively correct place at the beginning of the minimization. In other words, we do not observe cations migrating from the solution to close proximity of the G-rich
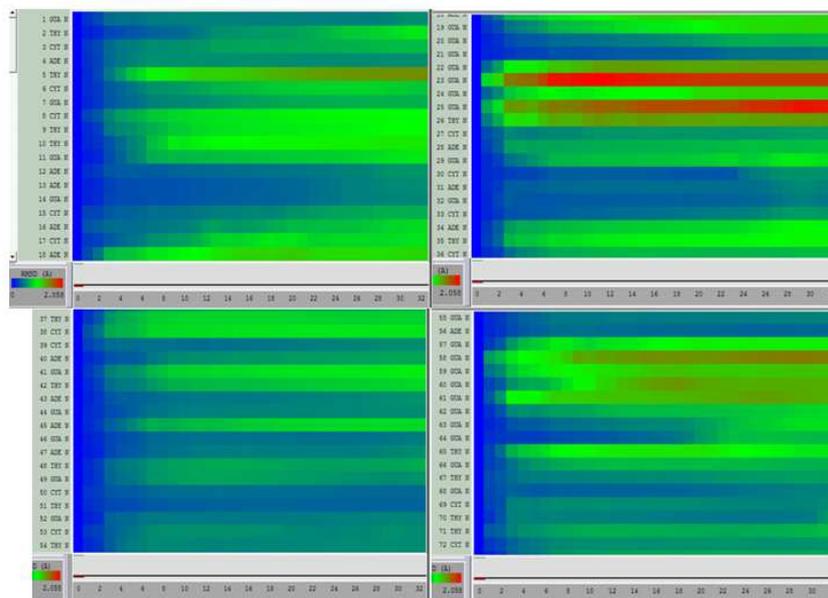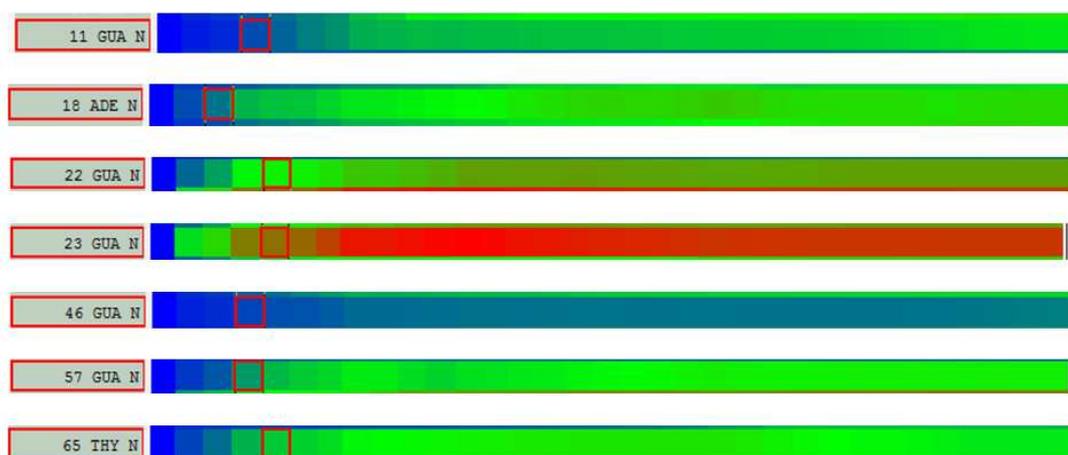
Figure 5: (a) RMSD for the SQ1 structure after minimization (geometry optimization). Each frame corresponds to 250 steps of minimization. (b) RMSD for selected bases from SQ1. Gua 11, guanine base in the interior of the double stranded region; Ade 18, base on the transition area between double stranded B - DNA and the zipper-like region; Gua 22 and Gua 23 bases in the center of the guanine rich zone. Gua 46, randomly selected base in the double stranded B-DNA region. Gua 57, guanine in the transition zone. Thymine 65 (THY 65), the first DNA base in the transition region.

region in the model prepared without the cations manually positioned close to these G bases.

Conclusions

The structural study of guanine risch DNA has implications and importance inboth basic and applied biology of cancer, new therapies, and possibly nano-electronics. G-rich sequences are widely distributed in the human genome complying roles for the regulation of gene expression and maintenance of chromosomes; offering unique targets for the future development of anti-cancer drugs [22]. In particular, the recent advances in biochemistry have allowed the construction of various types

of DNA to be selectively used as new therapeutic agents or as targets for these agents [22, 23]. In public databases, we have access to detailed structural models that can be used for discovery / drug design based on these structures. This is why DNA is such an important bio-molecule in the field of nanotechnology specially because the predictability of its interactions.
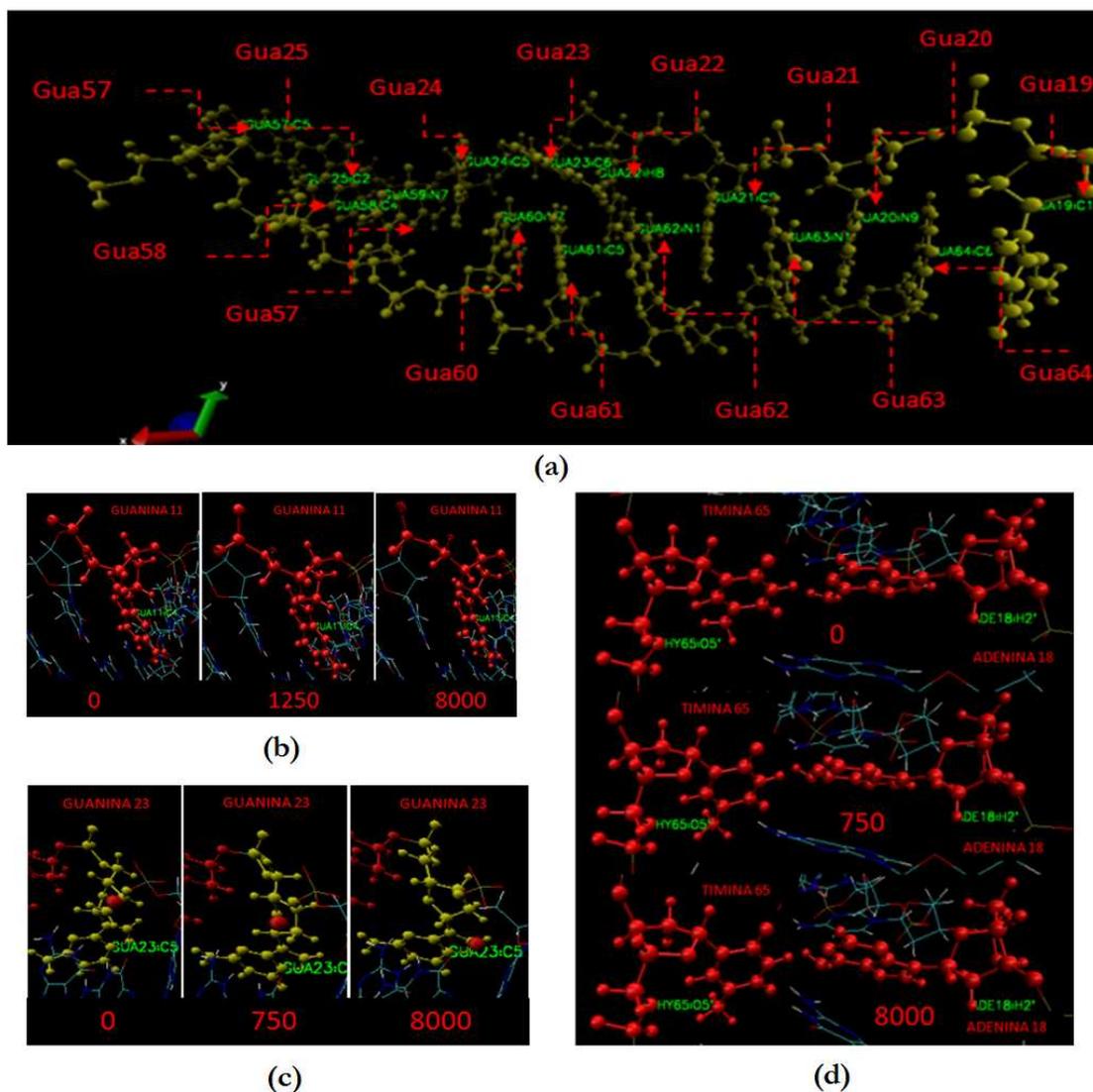
Acknowledgment

Figure 6: SQ1 optimized geometry corresponding to the G-rich region and the duplex region. (a) guanine-rich zone of the central part. (b) Guanine chosen randomly from the duplex region. (c) Guanine 23 located in the center of the guanine-rich region. (d) Base pair immediately before the region of guanines.

Group for the Study of Experimental and Theoretical Nanosystems) without its help and support this work would not have been possible. We thank F. Javier Torres, director of the Computational Chemistry and Theoretical USFQ (QCT-USFQ) for logistical support and project review. To Veronika A. Szalai with whom we are studying the assembly properties of these sequences.

References

[1] Seeman, N. 2007. "An Overview of Sructural DNA Nanotechnology". *Molecular Biotechnology.*. 37, 246–257.

[2] Park, K., Park, W.H., Sik, J., Dong, S., and Han, K. 2009. "Biomedical Polymer Nanofibers for Emerging Technology". *NanoScience in Biomedicine*. 21–42.

[3] Sadhasivam, S. and Yun, K. 2010. "DNA Self-Assembly: Prospectus and its Future Application". *Journal of Materials Science*. 45, 2543–2552.

[4] Shepard, W., Cruse, B., Fourme, R., de la Fortelle, E., and Prange, T. 1998. "A Zipper-Like Duplex in DNA:

the Crystal Structure of d(GCGAAAGCT) at 2.1 å Resolution". *Structure*. 6, 849–861.

[5] Balaeff, A., Craig, S., and D., B. 2011. "B-DNA to Zip-DNA: Simulating a DNA Transition to a Novel Structure with Enhanced Charge-Transport Characteristics". *The Journal of Physical Chemistry A*. 115, 9377–9391.

[6] Flores, O., Rendón, J.L., Martínez, F., Guerra, G., Sierra, E., and Pardo, P.J. 2008. "Las Herramientas del Modelado Molecular".

[7] Chou, S., Chin, K., and Wang, A. 2003. "Unusual DNA Duplex and Hairpin Motifs". *Nucleic Acids Research*. 31, 2461–2474.

[8] Lu, X. and Olson, W. 2008. "3dna: a Versatile, Integrated Software System for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic-Acid Structures". *Nature Protocols*. 3, 1213–1227.

[9] Kondo, J., Ciengshin, T., Juan, E. C.M., Sato, Y., Mitomi, K., Shimizu, S., and Takénaka, A. 2006. "Crystal Structure of d(gcGXGAgc) with X = G: a Mutation

at X is Possible to Occur in a Base-Intercalated Duplex for Multiplex Formation". *Nucleosides, Nucleotides and Nucleic Acids.* 25, 693–704.

[10] Špačková, N.A., Berger, I., and Šponer, J. 2000. "Nanosecond Molecular Dynamics of Zipper-Like DNA Duplex Structures Containing Sheared G•A Mismatch Pairs". *Journal of the American Chemical Society.* 122, 7564–7572.

[11] Venczel, E.A. and Sen, D. 1996. "Synapsable DNA". *J. Mol. Biol..* 257, 219–224.

[12] Burge, S., Parkinson, G., Hazel, P., Todd, A., and Neidle, S. 2006. "DNA: Sequence, Topology, and Structure". *Nucleic Acids Research.* 34, 5402–5415.

[13] Hud, N. V., Smith, F. W., Anet, F. A. L., and Feigon, J. 1996. "The Selectivity for K^+ versus Na^+ in DNA Quadruplexes is Dominated by Relative Free Energies of Hydration: a Thermodynamic Analysis by ^{1}HNMR". *Journal of Biochemistry.* 35, 15383–15390.

[14] Méndez, M. and Montero, A. 2011. *Avances en Ciencias e Ingenierías.* 3, A35–45.

[15] Endres, R., Cox, D., and Singh, R. 2004. "Colloquium: The Quest for High Conductance DNA". *Reviews of Modern Physics.* 76, 195.

[16] vonKiedrowski, G., Naumann, K., Pankau, W., Reimold, M., and M., R. 2003. "Toward Replicatable, Multifunctional, Nanoscaffolded Machines. A Chemical Manifesto ". *Pure Applied Chemestry.* 75, 609–619.

[17] Ongaro, A., Griffin, F., Beecher, P., Nagle, L., Lacopino, D., Quinn, A., et.al. 2005. "DNA-templated Assembly of Conducting Gold Nanowires Between Gold Electrodes on a Silicon Oxide Substrate". *Chemistry of Materials.* 17, 1959–64.

[18] Rohs, R., Jin, X., West, S .M. Joshi, R., Honig, B., and Mann, R. 2010. "Origins of Specificity in Protein-DNA Recognition". *Annual Reviews of Biochemistry.* 79, 233–269.

[19] Gellert, M., Lipsett, M.N., and Davies, D.R. 1962. *Proccedings of the National Academy of Science U.S.A.* 48, 2013–2018.

[20] Savelyev, A. and Papoian, G.A. 2007. "Inter-DNA electrostatics from Explicit Solvent Molecular Dynamics Simulations". *Journal of the American Chemical Society* 129, 6060–6061.

[21] Lee, J., Latimer, L., and Reid, R. 1993. "A Cooperative Conformational Change in Duplex DNA Induced by Zn^{2+} and other Divalent Metal Ions". *Biochemistry and Cell Biology.* 71, 162–8.

[22] Read, M.A., Wood, A.A., Harrison, J.R., Gowan, S.M., Kelland, L.R., Dosanjh, H.S., and Neidle, S. 1999. "Molecular Modeling Studies on G-Quadruplex Complexes of Telomerase Inhibitors: Structure-Activity Relationships". *Journal of Medical Chemistry.* 42, 4538–4546.

[23] Kelley, S., Boroda, S., Musier-Forsyth, K., and Kankia, B. I. 2011. "HIV-Integrase Aptamer Folds into a Parallel Quadruplex: A Thermodynamic Study". *Biophysical Chemistry.* 155, 82–88.