The effect of carbon nanotube on the structure of H–NS protein DNA complex: molecular dynamics approach

Najmeh Mahdavipour, Mohammad Reza Bozorgmehr*, Mohammad Momen-Heravi

Department of Chemistry, Mashhad Branch, Islamic Azad University, Mashhad, Iran

*bozorgmehr@mshdiau.ac.ir, mr_bozorgmehr@yahoo.com

Most of the experimental biophysical and biochemical observations of proteins are in dilute solutions, while inside the cell is a crowded environment.The effect of crowding on the structure and activity of biomolecules is not completely clear. In this work, molecular dynamics simulation was used to study the effect of single walled carbon nanotube (SWCNT) on the H–NS protein in the presence and absence of double-stranded nucleic acid. The values of root mean square deviation (RMSD) and its distribution, radius of gyration (Rg) and its distribution and root mean square fluctuation (RMSF) were calculated. Changes in the secondary structure of the H–NS were also calculated. The contributions of each residue of H–NS in free energy of binding between H–NS and DNA were calculated. The results indicate that the SWCNT unfolds the structure of the H–NS. In terms of contribution of residues in secondary structures, in the presence of a SWCNT, the sheet secondary structure of the H–NS changes more than helices secondary structure. In the triple system, which includes H–NS, SWCNT and DNA; Ala-1, Arg-3, Lys-6, Lys-17, Arg-24, Lys-30, Lys-31, Lys-38 and Lys-46 residues have a favorable effect on the interaction of the H–NS with the DNA.

Keywords: crowding, MMPBSA, secondary structure, contact map.

Received: 21 October 2019

Revised: 23 November 2019

1. Introduction

In cellular environments, a large number of compounds, such as proteins, saccharides, lipids, DNA, RNA, ions, etc., coexist [1, 2]. The concentration range of biomolecules in the cell is between 40 and 500 gL^{-1} [3, 4]. This concentration range is equivalent to occupying 20 to 40 percent of the cell volume. Therefore, inside the cell is a crowded environment. However, most biochemical *in vitro* studies are carried out in dilute solution [5]. In the case of parameters such as temperature, pressure, and pH, the change of parameters can be such as to reflect the physiological behavior of the cell environment, but this is not true about the effect of crowding. The effect of crowding on the structure and activity of biomolecules is not completely clear. On the other hand, the understanding of this effect leads to a better understanding of the behavior of biomolecules in the cell's environment. In order to mimic molecular crowding *in vitro*, various materials are used in the laboratory. Alcohols [6], polymers [7], and carbon nanotubes [8] are among the materials used in this field. Substances used as a crowding agent should have features such as chemically inert, high solubility in water, and varied in size. Except for high solubility in water, carbon nanotubes are suitable for other criteria. Carbon nanotubes are biocompatible because they have carbon; on the other hand, they can have a wide variety in terms of length and diameter. Also, single-walled carbon nanotubes (SWCNT) have great applications in drug delivery [9] and providing biosensors [10] due to their unique structural features.

H–NS are DNA-binding proteins which can cause condense DNA *invivo* and *in vitro* [11]; it does this as eukaryotic histones. The H–NS protein has two domains, which are interacting with the DNA in the C-terminal domain. Protein H–NS has been shown to bind to DNA with 47 residues at its C-terminus, and digest the protein during this process [11]. The binding domain of H–NS protein has a unique three-dimensional structure that is not similar to other DNA-binding proteins. The DNA-binding domain of H–NS protein has two beta-sheetsand two-alphahelices. Due to the presence of six lysine residues and two arginine residues in the binding domain of H–NS, this region of the protein has a positive charge. H–NS has a binding affinity with all types of DNA, but its binding affinity to double-stranded DNA is more pronounced [12]. Two mechanisms for the binding of H–NS to DNA have been proposed [13]: (i) via specific binding, in which H–NS binds to an AT-rich region of DNA; (ii) non-specific binding, in which the H–NS can bind over the whole DNA.

Although being crowded *in vivo*, biomolecules maintain their structure and activity; however, crowing under *in vitro* conditions causes changes in the structure and function of the biomolecules. How the structure of proteins in the presence of crowding factors is changed is not fully understood. Here, the structure of H–NS protein and its interaction with double-stranded DNA using molecular dynamics simulation is investigated. This is done in the presence and absence of arm-chair single-wall carbon nanotubes.

2. Molecular dynamics simulation details

Four different simulation boxes were designed. In the first box, the binding domain of H–NS protein with pdb code 1HNS in the protein data bank was placed in the center of the simulation box (denoted as *protein* system). In the second box,the H–NS was placed along with the four arm-chair SWCNTs (5, 5) (denoted as *protein-SWCNT* system). The H–NS was placed in the center of the box and the SWCNTs were randomly placed. In the third box,the H–NS and the DNA were placed (denoted as *protein-DNA* system) (see Table 1). The DNA structure was taken from the protein databank with code 4QJU. In this structure, the protein is also existed with nucleic acid that the protein is removed using the PyMol software [14]. In the fourth box, the H–NS, DNA and four SWCNTs were placed (denoted as *protein-SWCNT-DNA* system). All designated boxes are filled with water type TIP3P [15]. Sodium and chloride ions were used to neutralize the design systems.

TABLE 1. Overview of studied system and simulation details

System	Number of water molecules	Number of NA ion(s)	Number of Cl ion(s)
Protein	3891	0	1
Protein-SWCNT	5064	0	1
Protein-DNA	15895	39	0
Protein-SWCNT-DNA system	15664	39	0

All molecular dynamics simulations were performed with Gromacs package version 5.1.25.2.1 [16] and OPLS-AA/Lall atomforce field. Because the force field parameters for SWCNT are not present in the default version of the Gromacs software,the structure of this compound was optimized by using of B3LYP density functional method with the basis function of 6-31G*. To control the optimization, frequency calculations were performed and virtual frequencies were not observed. All *ab initio* calculations were done by GAMESS software package [17]. To eliminate the primary kinetic energy in each of the simulation boxes and inappropriate contacts between the atoms, energy minimization was done using steepest descent method [18]. Each simulation box achieved two-stage equilibrium in NVT and NPT ensemble. At this stage,the time of equilibration was considered 5 ns with time step 0.002 ps. Finally, molecular dynamics was performed by solving equations of motion for 100 ns with the 2 fs time step. The PME algorithm was used to calculate electrostatic interactions [19]. LINCS algorithm [20] was employed to fix the chemical bonds between the atoms of the protein and SETTLE algorithm [21] in the case of water molecules. To fix a constant temperature (298 K) and pressure (1 bar) during the simulations, systems components were coupled with V-rescaleand Nose-Hoover thermostat [22] respectively, in each of equilibration steps and molecular dynamics simulations.

3. Results and discussion

One way to examine the stability of a structure in molecular dynamics simulation is to calculate root mean square deviation (RMSD). The RMSD is defined as:

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^{N} m_i \left\| r_i(t_1) - r_i(t_2) \right\| \right]^{1/2},$$ (1)

where r_i is the atomic position at time t and $M = \sum_{i=1}^{N} m_i$. The Cα-RMSD values of the H–NS protein were calculated in the designated systems. The result is shown in Fig/ A1 of the Appendix. The stability of H–NS structure in systems *protein* and *protein-SWCNT-DNA* is similar,while H–NS in system *protein-DNA* has the lowest stability. For a better comparison between structural stability in different systems, the probability distribution of the RMSD value was calculated. The result is shown in Fig. 1.

According to the figure, in the *protein-SWCNT-DNA* system with the most crowding H–NS is similar to H–NS in system *protein*. Also, SWCNT alone and nucleic acid alone cause instability of the H–NS structure. Another factor that can determine the activity and stability of the protein structure is compression of the conformation [23]. We used the radius of gyration (*Rg*)to indicate the compression of the protein structure. The *Rg* value of protein is calculated
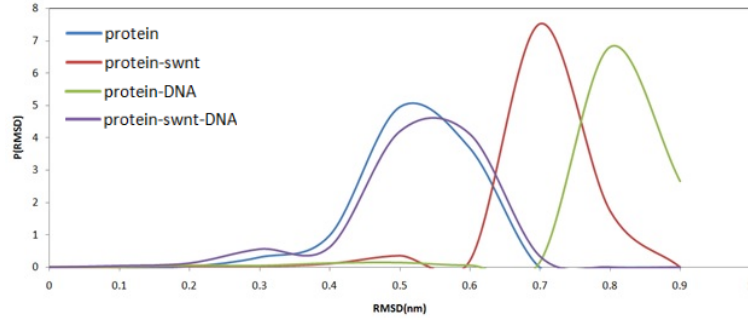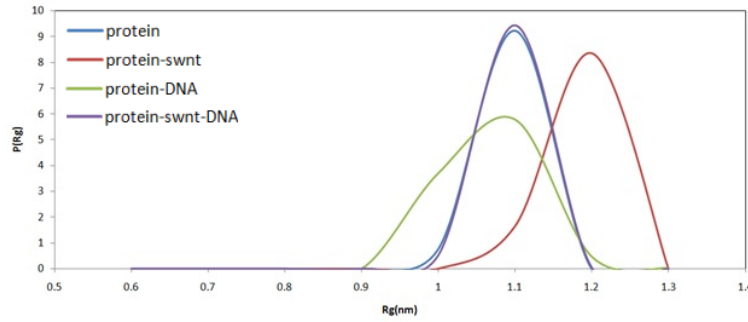
FIG. 1. The probability distribution of the RMSD in systems: *protein*, *protein-SWCNT*, *protein-DNA* and *protein-SWCNT-DNA*

using following relation:

$$R_g = \left(\frac{\sum_i \left\| r_i \right\|^2 m_i}{\sum_i m_i} \right)^{\frac{1}{2}}, \tag{2}$$

where m_i is the atomic mass of i and r_i the atomic position of i relative to the center of the molecule [24]. The R_g of H–NS were calculated in the designed systems. The result is shown in Fig. A2 of the Appendix. The probability distribution of the R_g value was calculated, also. The result is shown in Fig. 2.

FIG. 2. The probability distribution of the R_g in systems: *protein*, *protein-SWCNT*, *protein-DNA* and *protein-SWCNT-DNA*

Regarding the figure, it can be seen that the SWCNT has caused the H–NS structure to unfold. Since the SWCNT has a hydrophobic nature, it is expected to change the structure of the protein by changing the hydrophobic interactions. In other systems, the average of R_g is almost equal. The similarity between H–NS in systems *protein* and *protein-SWCNT-DNA* is significant. The root mean square fluctuation (RMSF) values of H–NS sequences are calculated in different systems and shown in Fig. 3.

RMSF is a measure of the flexibility of protein's residues [25]. It is observed at the N-terminal of the H–NS, where the fluctuations are greater, especially in the case of the *protein-SWCNT-DNA* system. Also, in the middle region of the H–NS, the Arg-24 residue in the *protein-SWCNT-DNA* system has a high flexibility. Given the fact that this residue is in the random coil region of H–NS, and its side chain is larger than any other residue, this result is logical. To further examine the changes in the residues in the secondary structure of the H–NS in different studied systems, the secondary structure of the H–NS was calculated in each of the designed systems. The results are shown in Fig. 4.

The secondary structure of the H–NS, along with solvent accessible surface area for its residue, was obtained from polyview-2d software [26]. To determine the secondary structures shown in Fig. 4, a sample structure from molecular trajectory is required. The free energy landscape analysis (FEL) method was used for sampling [27].

There are three steps in the free energy analysis method: first, calculate the root-mean-square-deviation (RMSD) and the H–NS's radius of gyration (R_g); second, calculate the probability of the presence of H–NS's conformations in each of the values calculated for the RMSD and R_g; third, calculate the free energy based on the probability values calculated in the second step. The results of the FEL analysis are reported in two-dimensional form in Fig. A3 through A6 and in 3D in Figs. A7 to A10 in the Appendix. Regarding these figures, a minimum free energy region was
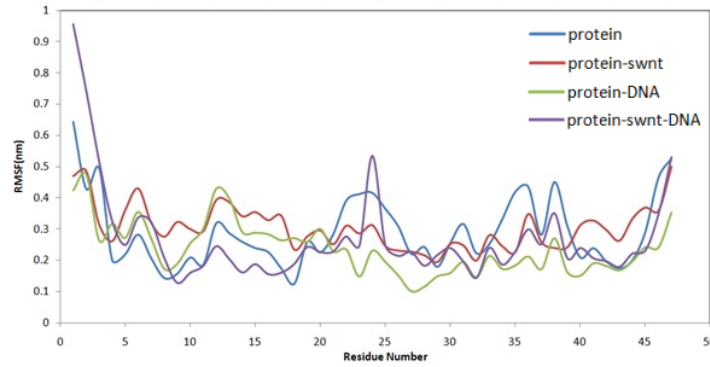
FIG. 3. Average RMSF values for each residue of H–NS during simulation in each system

FIG. 4. The secondary structure of the H–NS, along with the solvent accessible surface area for its residues.

Helix; β strand;   coil;   Dark: completely buried, bright: completely exposed

sampled. Fig. 4A is obtained from the free energy sampling. As can be seen in Fig. 4A, in the secondary structure of H–NS in systems *protein* and *protein-SWCNT-DNA* there are two helices and two sheets. In the secondary structure of H–NS, in systems *protein-SWCNT* and *protein-DNA*, the two sheets structures have disappeared. The first sheet consists of Ser-8, Tyr-9 and Val-10 residues. Considering the aromatic ring of Tyr-9 and the structure of two other amino acids, this sheet has a hydrophobic nature. The second sheet is composed of Thr-16, Lys-17 and Thr-18 amino acids. Therefore, the second sheet that is deleted in systems *protein-SWCNT* and *protein-DNA* has a polar nature. In system *protein-SWCNT-DNA*, the first size of first helix is larger than the first helix in the system *protein*. On the other hand, the size of the second helix in both systems is approximately equal. It is also observed that in system *protein-SWCNT-DNA*, the first sheet has been folded into the H–NS interior, while in the *protein* system the solvent accessible surface area of first sheet is increased. The contact map was used to check the change in the tertiary structure of the H–NS. Fig. 5 shows the contact map and the contact difference maps of the structures of the H–NS in *designed* system.

In Fig. 5, in the lower triangle of the diagram the black dots indicate the common contacts and the pink dots show the contacts that are present in the H–NS in *protein* system, but absent in the H–NS in *protein-SWCNT* system. The green dots indicate the contacts that are present in the H–NS in *protein-SWCNT* system, but absent in the H–NS in *protein* system. In the upper triangle of the diagram, the differences between the structures are indicated by the intensity of the red and blue colors. The regions in blue color show the contacts that are not altered and those in the red color show the difference between the two structures. The greater the number of red dots in Fig. 5 shows that the SWCNT alone have had a greater effect on the tertiary structure of H–NS. On the other hand, the distribution pattern of red spots is different in Fig. 5, which shows how the SWCNT effects on the tertiary structure differ in the presence and absence of DNA. The Kollman et al. method was used [28, 29] used to determine the contribution of each residue of H–NS in free energy of binding between H–NS and DNA. In this method, the free energy is obtained by:

$$G = E_{bnd} + E_{el} + E_{vdW} + G_{pol} + G_{np} - TS, \tag{3}$$

where G_{pol} and G_{np} are polar and nonpolar solvation free energies, which were obtained from the generalized Born and solvent accessible surface methods, respectively and E_{bnd}, E_{el} and E_{vdW} are the MM energies (bonding, bending, and dihedral), electrostatic energy and van der Waals interactions, respectively. The last term in Eq. 3, in which S is entropy and T is temperature, is obtained from the normal mode (NM) analysis. The contribution of each of the H–NS's residues in the free energy of binding of H–NS to the DNA in designed system is shown in Fig. 6.

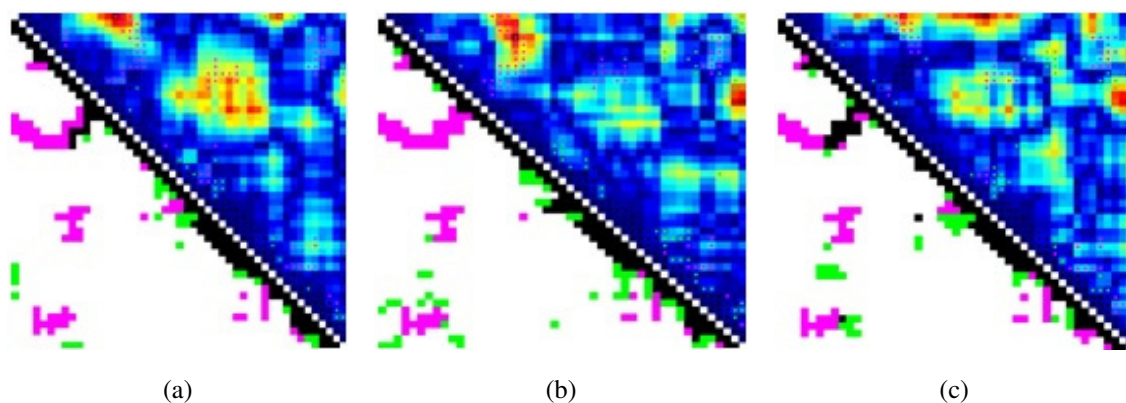(a)   (b)   (c)

FIG. 5. The contact map (lower triangle) and the contact difference maps (upper triangle) of the A) structures of the H–NS in *protein* system and H–NS in *protein-SWCNT* system B) the structures of the H–NS in *protein* system and H–NS in *protein-DNA* system C) the structures of the H–NS in *protein* system and H–NS in *protein-SWCNT-DNA* system. The colors used are described in the text
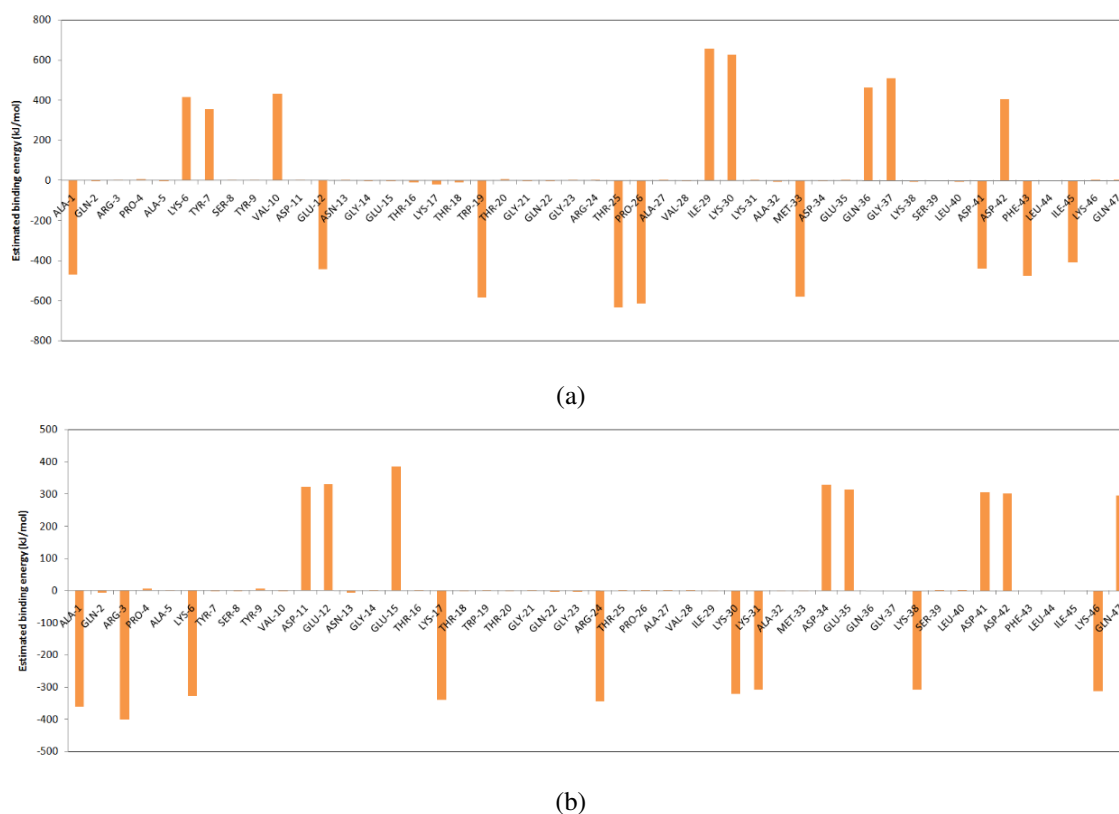

(a)

(b)

FIG. 6. The contribution of each of the H–NS's residues in the free energy of binding of H–NS to the DNA in A) *protein-DNA* B)*protein-SWCNT-DNA*

Comparison of A and B in Fig. 6 shows that in both systems eight residues have positive free energy (unfavorable effect) in H–NS binding to DNA, and nine amino acids have negative free energy (favorable effect). In system *protein-DNA*, residues that have an unfavorable effect on H–NS binding to the DNA are Lys-6, Tyr-7, Val-10, Ile-29, Lys-30, Gln-36, Gly-37 and Asp42. Also, in this system, residues that have a favorable effect on H–NS binding to the DNA are: Ala-1, Glu-12, Trp-19, Thr-25, Pro-26, Met-38, Asp-41, Phe-43 and Ile-45. In system *protein-SWCNT-DNA*, residues that have an unfavorable effect on H–NS binding to the DNA are Asp-11, Glu-12, Glu-15, Asp-34, Glu-35, Asp-41, Asp-42 and Gln-47. In this system, residues that have favorable effect onH–NS binding to the DNA are Ala-1, Arg-3, Lys-6, Lys-17, Arg-24, Lys-30, Lys-31, Lys-38 and Lys-46. It is observed that the SWCNT has affected

the interaction of H–NS with DNA by changing the position of charged residues. In the absence of a SWCNT, it also seems that the H–NS interacts more strongly with hydrophobic interactions with DNA, while in the presence of SWCNT the H–NS interacts with the DNA by electrostatic interactions.

4. Conclusion

The effect of single walled arm-chair carbon nanotube (5, 5) on H–NS protein was studied by molecular dynamics simulation. The calculations were carried out in the presence and absence of double-stranded nucleic acid. In the secondary structure of the H–NS, the beta sheets of the protein were influenced by the SWCNT more than their helices. In a triple system that contains H–NS, DNA and SWCNT, and is more crowded, the structure of the H–NS is more similar to that of the dilute solution. The interactions between protein and nucleic acid and the binding energies of the residues involved were obtained. The results indicate that the nanotube has altered the interaction between protein and nucleic acid. The results are consistent with experimental observations [30].

References

[1] Ellis R.J. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current opinion in structural biology*, 2001, **11** (1), P. 114–119.

[2] Miyoshi D., Sugimoto N. Molecular crowding effects on structure and stability of DNA. *Biochimie*, 2008, **90** (7), P. 1040–1051.

[3] Srere P.A. Protein crystals as a model for mitochondrial matrix proteins. *Trends in Biochemical Sciences*, 1981, **6**, P. 4–7.

[4] Zimmerman S.B., Minton A.P. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual review of biophysics and biomolecular structure*, 1993, **22** (1), P. 27–65.

[5] Ellis R.J. Macromolecular crowding: obvious but underappreciated. *Trends in biochemical sciences*, 2001, **26** (10), P. 597–604.

[6] Doghaei A.V., Housaindokht M.R., Bozorgmehr M. Molecular crowding effects on conformation and stability of G-quadruplex DNA structure: Insights from molecular dynamics simulation. *Journal of theoretical biology*, 2015, **364**, P. 103–112.

[7] Jiao M., et al. Attractive protein-polymer interactions markedly alter the effect of macromolecular crowding on protein association equilibria. *Biophysical journal*, 2010, **99** (3), P. 914–923.

[8] Xu M., et al. Alignment control of carbon nanotube forest from random to nearly perfectly aligned by utilizing the crowding effect. *ACS Nano*, 2012, **6** (7), P. 5837–5844.

[9] Bianco A., Kostarelos K., Prato M. Applications of carbon nanotubes in drug delivery. *Current opinion in chemical biology*, 2005, **9** (6), P. 674–679.

[10] Tlmaciu C.-M., Morris M.C. Carbon nanotube biosensors. *Frontiers in chemistry*, 2015, **3**, P. 59.

[11] Atlung T., Ingmer H. HNS: a modulator of environmentally regulated gene expression. *Molecular microbiology*, 1997, **24** (1), P. 7–17.

[12] Falconl M., et al. Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15kD Escherichia coli DNA binding protein HNS. *Molecular microbiology*, 1988, **2** (3), P. 323–329.

[13] Tupper A.E., et al. The chromatinassociated protein HNS alters DNA topology in vitro. *The EMBO journal*, 1994, **13** (1), P. 258–268.

[14] DeLano W., The PyMOL Molecular Graphics System. DeLano Scientific; Palo Alto, CA: 2002. URL: http://www.pymol.org.

[15] Price D.J., Brooks C.L. III. A modified TIP3P water potential for simulation with Ewald summation. *The Journal of chemical physics*, 2004, **121** (20), P. 10096–10103.

[16] Van Der Spoel D., et al. GROMACS: fast, flexible, and free. *Journal of computational chemistry*, 2005, **26** (16), P. 1701–1718.

[17] Schmidt M.W., et al. General atomic and molecular electronic structure system. *Journal of computational chemistry*, 1993, **14** (11), P. 1347–1363.

[18] Luenberger D.G., Ye Y. *Linear and nonlinear programming*, **2**, 1984, Springer.

[19] Essmann U., et al. A smooth particle mesh Ewald method. *The Journal of chemical physics*, 1995, **103** (19), P. 8577–8593.

[20] Hess B., et al. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 1997, **18** (12), P. 1463–1472.

[21] Miyamoto S., Kollman P.A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of computational chemistry*, 1992, **13** (8), P. 952–962.

[22] Bussi G., Donadio D., Parrinello M. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 2007, **126** (1), P. 014101.

[23] Ghaderi S., Bozorgmehr M.R., Morsali A. Structure study and predict the function of the diphtheria toxin in different pH levels (Acidic-Basic-Natural) using molecular dynamics simulations. *Entomology and Applied Science Letters*, 2017, **3** (4), P. 49–56.

[24] Honarparvar B., Skelton A.A. Molecular dynamics simulation and conformational analysis of some catalytically active peptides. *J. Mol. Model.*, 2015, **21** (4), P. 100.

[25] Housaindokht M.R., Bozorgmehr M.R., Monhemi H. Structural behavior of Candida antarctica lipase B in water and supercritical carbon dioxide: A molecular dynamic simulation study. *The Journal of Supercritical Fluids*, 2012, **63**, P. 180–186.

[26] Porollo A.A., Adamczak R., Meller J. POLYVIEW: a flexible visualization tool for structural and functional annotations of proteins. *Bioinformatics*, 2004, **20** (15), P. 2460–2462.

[27] Lei H., et al. Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. *Proceedings of the National Academy of Sciences*, 2007, **104** (12), P. 4925–4930.

[28] Kumari R., et al. g_mmpbsa–a GROMACS tool for high-throughput MM-PBSA calculations. *Journal of chemical information and modeling*, 2014, **54** (7), P. 1951–1962.

[29] Kollman P.A., et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*, 2000, **33** (12), P. 889–897.

[30] Nii D., et al. Selective binding of single-stranded DNA-binding proteins onto DNA molecules adsorbed on single-walled carbon nanotubes. *Colloids and Surfaces B: Biointerfaces*, 2014, **121**, P. 325–330.

Appendix

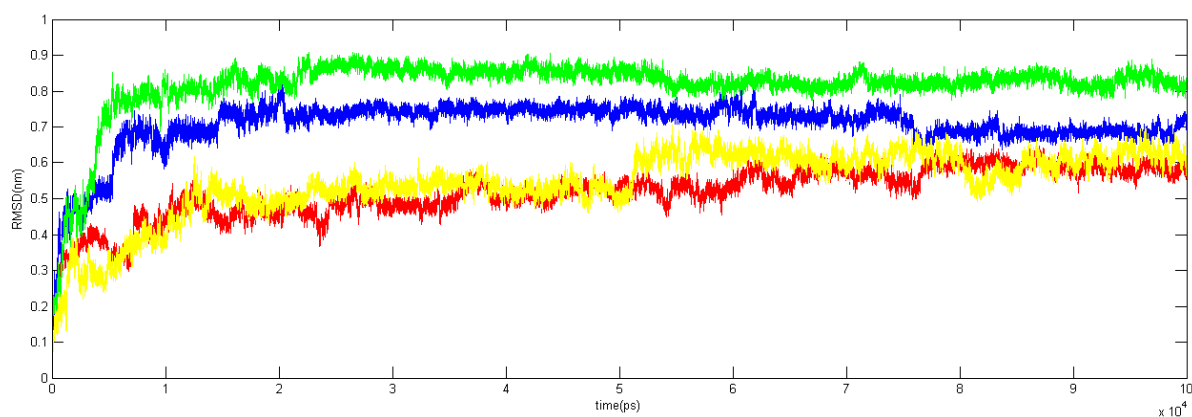

FIG. A1. The Cα-RMSD (nm) of H–NS vs. time (ps). Red: *protein*; blue: *protein-SWCNT*; green: *protein-DNA*; yellow: *protein-SWCNT-DNA*
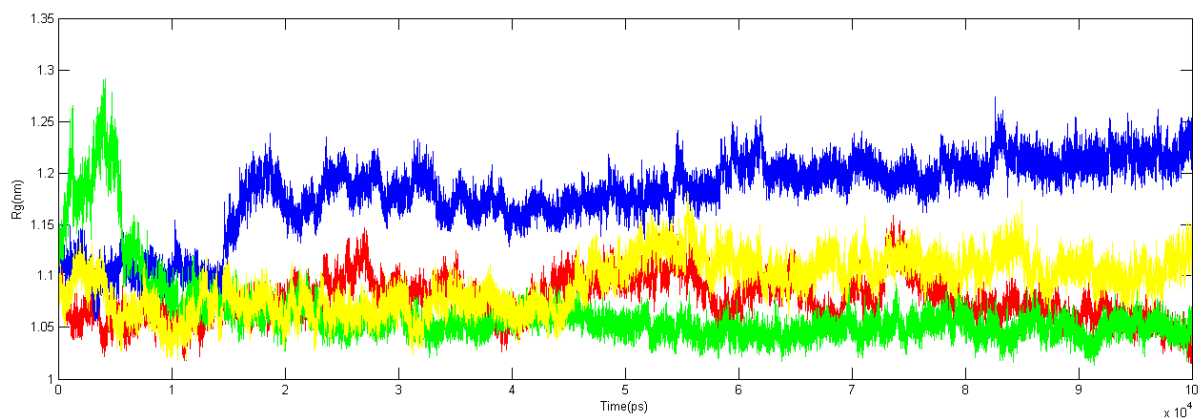

FIG. A2. The R_g (nm) of H–NS vs. time (ps). Red: *protein*; blue: *protein-SWCNT*; green: *protein-DNA*; yellow: *protein-SWCNT-DNA*
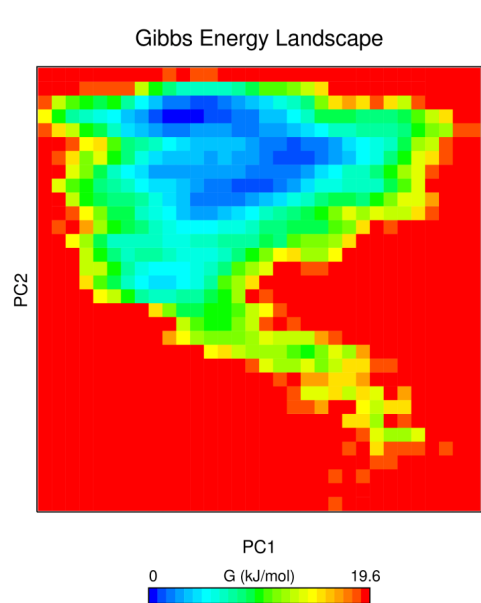
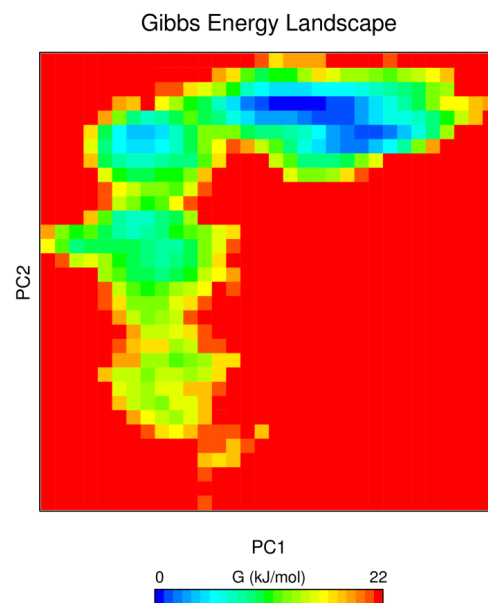FIG. A3. 2D Free Energy Landscape of the *protein* simulated system

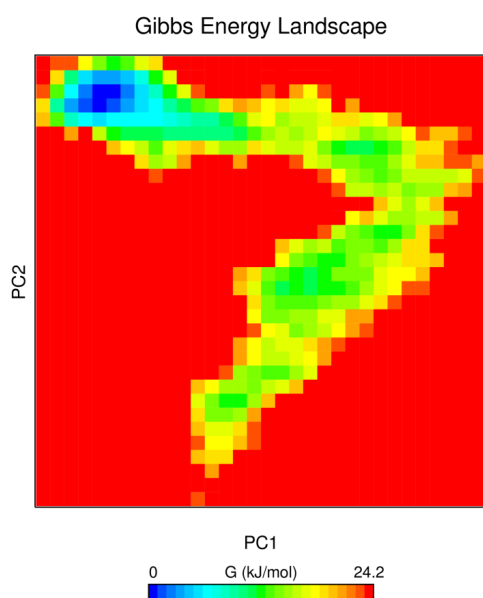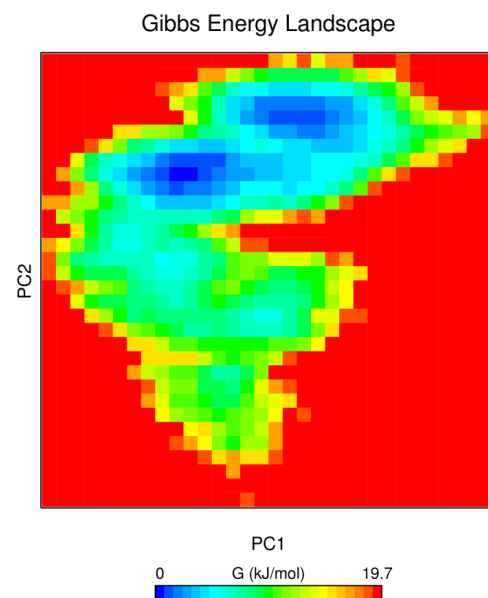FIG. A4. 2D Free Energy Landscape of the *protein-SWCNT* simulated system

FIG. A5. 2D Free Energy Landscape of the *protein-DNA* simulated system

FIG. A6. 2D Free Energy Landscape of the *protein-SWCNT-DNA* simulated system
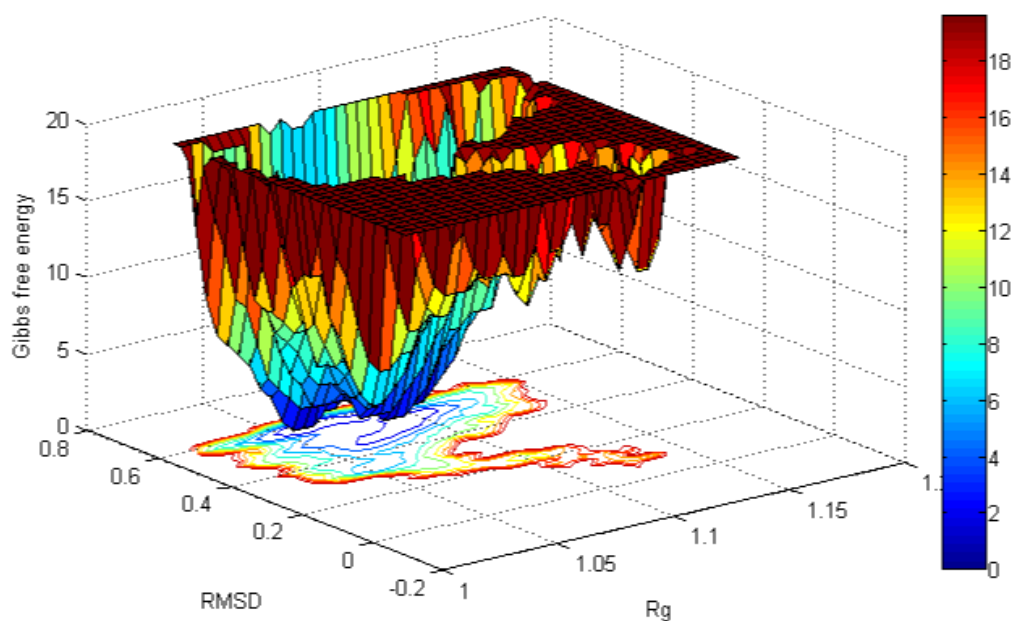
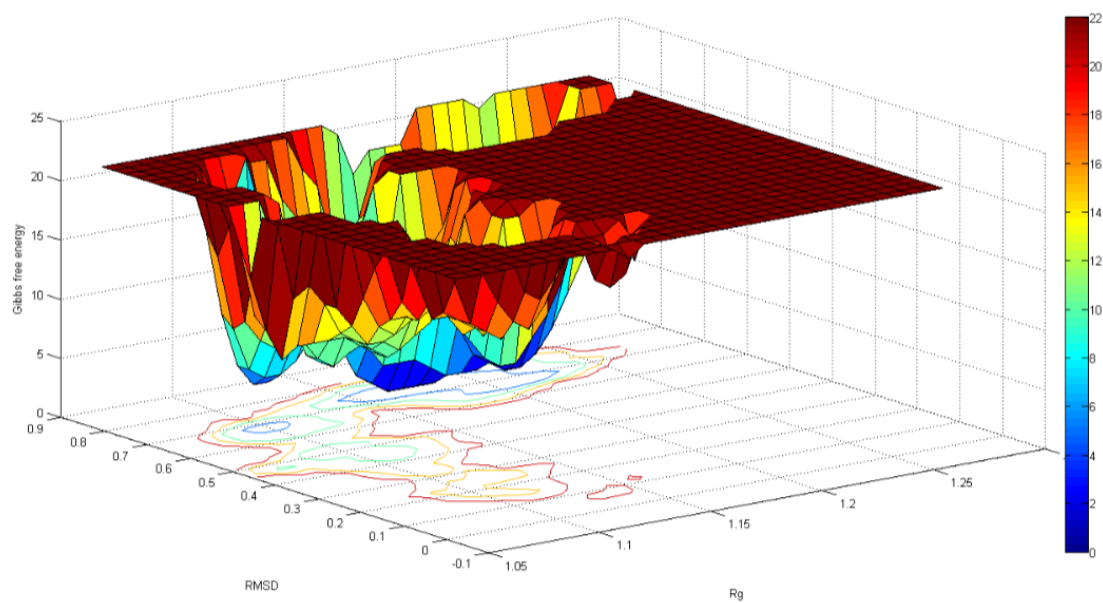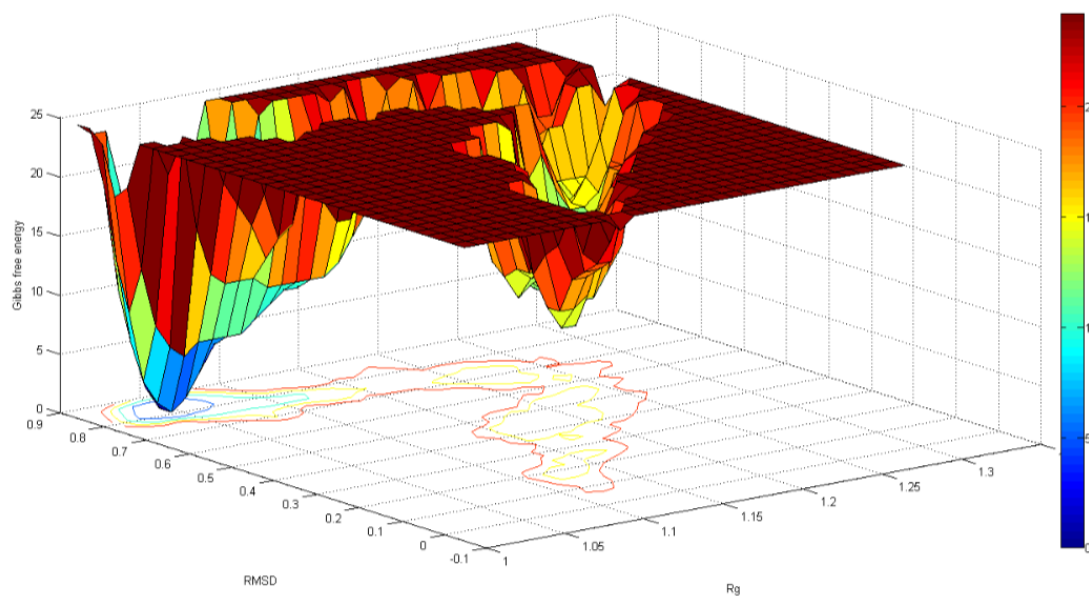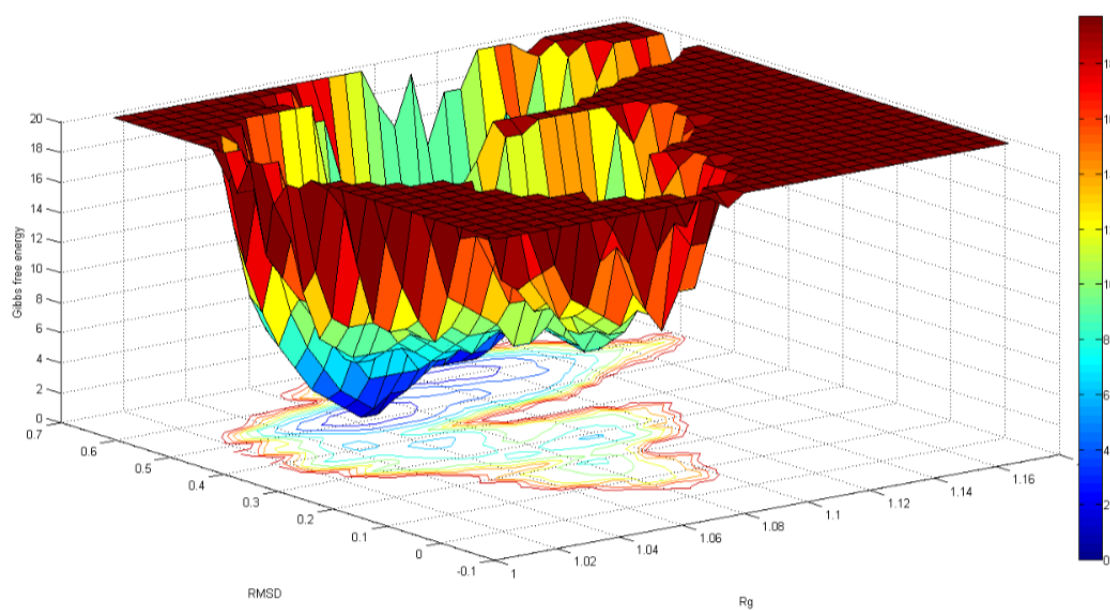FIG. A7. 3D Free Energy Landscape of the *protein* simulated system

FIG. A8. 3D Free Energy Landscape of the *protein-SWCNT* simulated system

FIG. A9. 3D Free Energy Landscape of the *protein-DNA* simulated system

FIG. A10. 3D Free Energy Landscape of the *protein-SWCNT-DNA* simulated system