Microscopic Studies on Biomolecular Complexes

Thesis submitted for the degree of

Doctor of Philosophy (Sc.)

in

Physics (Theoretical)

by

Sasthi Charan Mandal

Department of Physics

University of Calcutta

April, 2022

Dedicated to my Family

Acknowledgements

I express my heartfelt gratitude and sincere thanks to my supervisor Prof. Jaydeb Chakrabarti for his continuous support and invaluable guidance during my Ph.D. years. His enthusiasm and dedication towards research have always been a great inspiration to me. His hope and belief on me help me to pursue diverse problems simultaneously and also help me to develop myself as a good scientist.

I am grateful to my thesis and Ph.D. committee members Prof. Sanjoy Bandyopadhyay, IIT Kharagpur, Dr. Suman Chakrabarty and Prof. Gautam Gangopadhyay, S. N. Bose National Centre for Basic Sciences (SNBNCBS) for their helpful comments and suggestions. I am truly thankful to Dr. Manas Mondal, Dr. Lakshmi Maganti and Dr. Uttam Pal for their helpful suggestions. I thank to my seniors Piya Patra and Sutapa Dutta for their help and supports over the years of my Ph.D. It is a pleasure to acknowledge my fellow group members: Abhik Ghoushmoulik, Edwine Tendong, Rahul Karmakar, Anirban Pal, Suravi Pal, Aayatti Mallick Gupta, Kanika Kole and Avik Sasmal who helped me during different phases of my works in their own unique ways.

I would also express my gratitude towards Prof. Achintya Dhar, Dr. Moumita Patra and Dr. Pradipta Gupta Roy who not only enriched my knowledge but also inspired me to grow as a good human being. I thankfully acknowledge Raghunathpur College and IIT Kharagpur for providing me the perfect atmosphere and educational support.

I am lucky to have a large number of wonderful friends, and this journey would not have been possible without them. I specially thank Paban, Bapi, Pratap, Sayan Pal, Daisy, Abinash, Akash da, Suman da, Mithun da, Santanu, Sayan Routh, Parushottam, Dipanjan, Dibyendu, Ikbal, Samir, Shantanu, Saikat, Partha, Sumana, Imadul, Didhiti, Gesesew and Rajdeep. I am also thankful to the staffs of the mess and canteen of SNBNCBS for their service.

I sincerely thank to my parents Abani Mandal and Champa Mandal, and my sisters Jayanti Mandal and Kanan Mandal for always being there with me through all successes and failures. My journey would not have been possible without their love and support. I have been blessed to have such a wonderful and supportive didi Jayanti Mandal who has guided and supported me in all aspects of my life.

Finally, I thank S. N. Bose National Centre for Basic Sciences for providing me a wonderful research environment and facilities. I am thankful to the faculty members and staffs of this centre for their cooperation and help. I gratefully acknowledge CSIR for providing me fellowship.

Sasthi Charan Mandal

Department of Chemical, Biological and Macro-Molecular Sciences

S. N. Bose National Centre for Basic Sciences

Block JD, Sector III, Salt Lake,

Kolkata-700106, West Bengal, India.

Abstract

The thesis is based on studies of two types of biomolecular complexes: (1) complexes between biomolecules and (2) complexes of biomolecules with inorganic moieties. These complexes are important for their applications in gene cloning, protein expression, drug design, sensors and detection of diseases.

We consider an important class of biomolecular complexes, namely, restriction endonucleases (RE)-DNA. REs protect bacterial cells against bacteriophage infection by cleaving the viral DNA into fragments. The cleavage of a specific DNA strand by the REs occurs in presence of Mg^{2+}, but not in presence of Ca^{2+}, although the proteins bind to DNA in presence of both metal ions. In order to understand the metal ion specificity in EcoRV, we compare the fluctuations of microscopic degrees of freedom, like DNA base pair step parameters and dihedral angles of protein residues using all-atom Molecular Dynamics (MD) trajectories of the complexes. We calculate the changes in free energy and entropy due to differences in conformations of the biomolecules in Mg^{2+} ion bound EcoRV-DNA (Mg^{2+}-EcoRV-DNA) complex compared to those in Ca^{2+} bound EcoRV-DNA (Ca^{2+}-EcoRV-DNA) complex. We find that the base pairs in the cleavage region are disordered in Ca^{2+}-EcoRV-DNA compared to Mg^{2+}-EcoRV-DNA. One of the acidic residues ASP90, coordinating to the metal ion in the vicinity of the cleavage site, is conformationally destabilized and disordered.

Previous studies suggest that water plays an important role in the DNA cleavage activity by the REs. We extend consider another complex of RE, EcoRI with DNA to investigate cofactor the DNA sequence in DNA cleavage, emphasizing the role of water. We consider four complexes, EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA where the second base pair in the recognition sequence in Mg^{2+}-EcoRI-DNA is mutated. We observe that the Mg^{2+}-EcoRI-DNA complex shows the maximum conformational stability and

order in the cleavage region compared to the free components among all the complexes. The number of hydrogen bonds (HBs) with the surrounding water molecules around the scissile phosphate group is less compared to that for the other phosphate groups in all the complexes. This number at the cleavage phosphate group is smaller in Mg^{2+}-EcoRI-DNA and mutated complexes compared to those of the other complexes. Hence, the HB network in the scissile phosphate group is less strong in Mg^{2+}-EcoRI-DNA and mutated complexes. This indicates that the probability of attacking the scissile phosphate group by a water molecule is more in these complexes compared to the other complexes. Among these two complexes, however, the conformational stability is higher in the Mg^{2+}-EcoRI-DNA than Mg^{2+}-EcoRI-mtDNA. Thus, Mg^{2+}-EcoRI-DNA is having the maximum efficiency in cleavage activity.

Next, we consider a couple of nano-bio complexes. We investigate the interactions of ZnO nanoparticles (ZnONP) with ATP, ADP and AMP using density functional theory (DFT) based quantum chemical calculations. Interactions of ZnONP with ATP, ADP and AMP are important to understand how the metal oxide nanoparticles affect the cellular energy transfer reactions. We establish chelation of the metal oxide surface involving Zn-O bond through ZnONP → phosphate CT and Zn-N bond via adenine → ZnONP CT and additional O-H---O non-classical bonds with ZnO nano-cluster with phosphate groups in all the nano-bio complexes. DFT-based Raman spectra strongly support the presence of these bonds in these nano-bio complexes.

ZnO is extensively used in non-enzymatic glucose biosensors for the detection of glucose concentrations in human blood. To understand the binding mechanism of glucose with four common surfaces, $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ respectively of ZnO, and examining the most preferable surface for glucose adsorption in aqueous solution, we carry out molecular dynamics (MD) simulations enhanced by umbrella sampling. We observe that the layered water molecules formed above the surfaces hinder the closer approach of glucose to

the surfaces. The glucose density is maximum near the surface of $(10\bar{1}0)$ compared to the other surfaces. Potential of mean force (PMF) calculations show that glucose molecules on $(10\bar{1}0)$ surface shows strongest adsorption free energy. Thus, the interactions between glucose and the surfaces are observed to be highly specific to the surface.

Table of Contents

CHAPTER 1

Introduction

Studying biomolecular complexes is important for their applications in biomedical science and biotechnological areas.[1–5] Biomolecular complexes are of two major kinds: (i) complexes of biomolecules with other biomolecules and (ii) complexes of biomolecules with inorganic surfaces, more commonly known as nano-bio conjugates. Complexes between biomolecules such as protein with nucleic acids, protein with small molecules, nucleic acids with small molecules etc. perform almost all biological processes in the cell, such as DNA replication, transcription, translation, protein expression to name a few.[6,7] On the other hand, the nano-bio conjugates attracts a great deal of attention due to their applications in biosensors, detection of diseases etc.[8–10]

The complexes of proteins with nucleic acids such as restriction enzymes-nucleic acids are responsible for performing many cellular functions.[7,11,12] Restriction endonucleases or restriction enzymes (REs) are proteins that cleave DNA into fragments at or near specific recognition sites in order to protect cells against bacteriophage infection.[11–13] REs occur ubiquitously in bacteria, archaea and in viruses of certain unicellular algae. The host DNA is protected by a modification enzyme that methylates the host DNA in order to block the cleavage by REs. REs are categorized into four types, type I, II, III and IV, depending on their subunit composition, cofactor requirement and mode of action.[14] Type II REs are extraordinarily important for their applications in gene cloning and gene analysis. Type II REs have been studied in great detail. Furthermore, they are outstanding model systems for studying highly specific protein-DNA interactions and structure-function relationships. Type II REs are homodimers that recognize a palindromic sequence of 4-8 base pairs in length. They cleave the two strands of the DNA in presence of Mg^{2+}.

Biomolecules also form complexes with inorganic materials. The interface between inorganic materials like ZnO, TiO2, etc. and biomolecules show unique physicochemical properties that has applications in nanotechnology and biomedical areas.[15–18] The chemical composition, surface functionalization, shape and angle of curvature, porosity, and surface crystallinity of nanomaterials all play a significant role in the conjugation of nanoparticles and biomolecules.[19,20] The primary important factors at the nano-bio interface are the nanoparticle surface, the solid-liquid interface and the contact zone of the solid-liquid interface with biomolecules. Water plays a critical role when considering the interactions of inorganic materials with biomolecules. So, understanding the nano-bio interface requires a thorough investigation of the solid-liquid interface.

In this thesis, we consider a couple of DNA-Restriction enzyme complexes. EcoRV is one of the type II REs that binds to GATATC sequence of DNA both in presence of Mg^{2+} and Ca^{2+}.[12,21,22] However, EcoRV cleaves DNA at the central base pair (TA) only in presence of Mg^{2+} but not in presence of Ca^{2+}.[23–25] The microscopic picture of the cofactor sensitivity of EcoRV in cleaving the DNA is yet to be known. We investigate metal ion sensitivity of the catalytic activity to understand this difference using conformational thermodynamics by calculating changes in conformational free energy and entropy of conformational degrees of freedom, like base pair steps of DNA and dihedral angles of protein residues in presence of metal ions, Mg^{2+} and Ca^{2+}, using the equilibrated trajectories of all-atom Molecular Dynamics (MD) simulations (as detailed in Appendix I of Chapter 2) of the complexes. If the distribution of fluctuations of a conformational degree of freedom are known between two conformational states of the macromolecule in a given ensemble, one can estimate the conformational free energy changes from the Boltzmann weight and the entropy changes from Gibbs formula by comparing the distributions in two states.

We observe that the base pair steps in the recognition region are all destabilized and disordered in Ca^{2+} loaded EcoRV-DNA (Ca^{2+}-EcoRV-DNA) complex compared to those in Mg^{2+} loaded EcoRV-DNA (Mg^{2+}-EcoRV-DNA) complex. The cleavage base pair step gets maximum disordered in presence of Ca^{2+}. One of the catalytically active residues ASP90 also gets destabilized and disordered in Ca^{2+}-EcoRV-DNA. The distance fluctuations between CG atoms of two catalytically active residues, ASP74 and ASP90 reveals that these two residues come closer in Ca^{2+}-EcoRV-DNA compared to those in Mg^{2+}-EcoRV-DNA, leading to larger electrostatic repulsion between these two residues in presence of Ca^{2+}. So, the cleavage region gets stabilized and ordered and the metal ion is stable in the neighbourhood of the scissile phosphate group in Mg^{2+}-EcoRV-DNA complex compared to Ca^{2+}-EcoRV-DNA complex. Replacement of the metal ions type in these complexes makes the cleavage region unstable and disordered. This suggests that both the type of the metal ions and their locations are important for cofactor sensitivity for cleavage by EcoRV.

Another type II RE, EcoRI, recognizes GAATTC sequence and cuts the DNA between G and A in presence of Mg^{2+}. Like EcoRV, the metal ion specificity of DNA cleavage is unknown for EcoRI also. Water plays a crucial role in DNA recognition and cleavage by EcoRV. According to the experiment, when EcoRI binds to the cognate sequence (GAATTC), it releases 70 fewer water molecules than when it binds to a noncognate sequence (TAATTC).[26] Furthermore, 30 water molecules are bound in the cognate sequence during the cleavage event. Although there are several experimental studies on the role of water in DNA cleavage,[26–28] a microscopic understanding is still elusive.

Here, we consider four systems, EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA respectively. Mg^{2+}-EcoRI-mtDNA is the complex where we mutate the second base pair (A:T) in the recognition region by (G:C) in Mg^{2+}-EcoRI-DNA complex. We also consider their free components such as EcoRI, Mg^{2+}-EcoRI, Ca^{2+}-EcoRI and free

DNA respectively. Our calculation is divided into two parts: (1) Metal ion specificity of DNA cleavage using conformational thermodynamics and (2) role of water in DNA cleavage due to cofactor sensitivity and DNA sequence sensitivity. First, we calculate the conformational free energy and entropy changes of the base pair step parameters of DNA base pair steps in EcoRI-DNA as well as metal ion loaded complexes with respect to free DNA. We observe that the cleavage region gets stabilized and ordered only in presence of Mg^{2+} ion as in the case of EcoRV. In order to find out the role of water molecules in DNA cleavage due to cofactor sensitivity and DNA sequence sensitivity, we calculate the hydrogen bonds (HBs) per water molecule (*f)* around each phosphate group of DNA in the recognition region. We observe that *f* is minimum for the cleavage phosphate compared to all other phosphate groups in all the complexes. At the cleavage phosphate group, *f* is minimum in both Mg^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes compared to the other complexes. This indicates that the HB network is less strong around the phosphate group in these complexes and more available for nucleophilic attack for cleavage. However, more fluctuations in the protein active sites in Mg^{2+}-EcoRI-mtDNA makes it unfavourbale for cleavage compared to Mg^{2+}-EcoRI-DNA.

We study a couple of nano-bio complexes. ZnO nanoparticles are useful in biotechnological applications[29–32] as it is less toxic, biocompatible and biosafe. It is important to understand how these nanoparticles interact with biomolecules prevalent in cellular conditions. ATP is an energy carrier biomolecule that provides energy to the cells by the process of hydrolysis[33,34]. Hydrolysis of ATP forms ADP, and further hydrolysis produce AMP[35,36]. During these hydrolysis processes, energy is released that fuel the other cellular processes[33,35,37]. We perform quantum chemical calculations based on DFT on a nano-bio conjugate between the $Zn_{12}O_{12}$ cluster (ZnONP) and ATP, ADP and AMP. In such calculations the ground state of the system is obtained by variationally minimizing the ground state energy of a fully interacting many-electron system with respect to the electron density, given by the

squared modulus of the ground state wave function and the interaction between the electrons modelled via correlation terms[38]. The hybrid functional B3LYP[39] is a standard method for a good compromise between computational cost, coverage, and accuracy of results. We show from our quantum chemical calculations, the existence of a Zn-N(adenine) bond and multiple Zn-O(phosphate) bonds in the DFT-optimized geometries. In all three nano-bio complexes, phosphate groups form additional O-H---O non-classical bonds with the ZnO nano-cluster. Natural population analyses and Kohn-Sham MOs suggest that the adenine moiety forms a coordinate covalent bond (Zn-N) with ZnO cluster through adenine → ZnO charge transfer (CT), while phosphate groups from coordinate covalent bonds (Zn-O) through ZnO → phosphate CT in all nano-bio complexes, leading to better stability of biomolecule-ZnO complex with increasing of number of phosphate groups.

We also examine the interactions between glucose molecules and ZnO crystal surfaces. An abnormal level of glucose in blood can lead to many diseases including diabetes, cardiovascular diseases, metabolic syndrome to name a few.[40,41] Thus, it is important to develop sensitive and reliable glucose biosensors in order to detect glucose in human blood. Glucose biosensors are mainly enzymatic. Due to disadvantages, like high cost for purification and fabrication, intrinsic instability and complicated immobilization procedure of enzyme-based glucose biosensors, it is desirable to develop enzyme free glucose biosensors. Such design depends on understanding the interaction between glucose and the common ZnO surfaces, like $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$. We carry out molecular dynamics (MD) simulations enhanced by umbrella sampling to study glucose adsorption on four different surfaces in aqueous solutions. The hydrated layer above the surfaces hinders the approach of glucose to the surface. The glucose density of 0.478 gm/cc is found to be maximum near the surface of $(10\bar{1}0)$ compared to the other surfaces. Potential of mean force (PMF) calculations also shows that $(10\bar{1}0)$ surface has the strongest adsorption free energy.

Our study reveals microscopically the importance of metal ions and water in DNA recognition and cleavage by the REs, which has long been debated. The microscopic fluctuations of the conformational variables may be useful to understand other DNA-protein enzyme complexes and in designing genetic tools for targeted DNA manipulations. The investigation on interactions of ZnO nanoparticles with energy carrier biomolecules can help to understand how nanoparticles affect the cellular energy transfer reactions. Our study on nano-bio complexes can also be applied to understand the microscopic interactions at the nano-bio junctions. This may help in designing the sensitive, fast and reliable biosensors.

The thesis is organised as follows: Chapter 2 contains the details of metal ions specificity of DNA cleavage by EcoRV using conformational thermodynamics. The details on changes in the conformational thermodynamics and hydration of DNA and metal ions in DNA-RE complexes to understand the specific recognition and cleavage of DNA by EcoRI is discussed in Chapter 3. Chelations of ZnO nanoparticles by energy carrier biomolecules, ATP, ADP and AMP are reported in Chapter 4. The surface specific interactions of glucose with ZnO surfaces are discussed in Chapter 5.

1.1 List of publications

The thesis is based on the following publications:

1. Chapter 2:"Microscopic insight to specificity of metal ion cofactor in DNA cleavage by restriction endonuclease EcoRV", **Sasthi Charan Mandal**, Lakshmi Maganti, Manas Mondal, Jaydeb Chakrabarti, *Biopolymers* e23396 (2020).

2. Chapter 3: "Cofactor and sequence sensitivity in DNA cleavage by EcoRI", **Sasthi Charan Mandal**, Jaydeb Chakrabarti, *Manuscript in preparation*.

3. Chapter 4: "Quantum chemical studies on chelation in nano-bio conjugate between ZnO nanoparticle and cellular energy carrier molecules", Mausumi Ray,[⊥] **Sasthi Charan Mandal**,[⊥] Jaydeb Chakrabarti ([⊥] These authors contributed equally to this work), submitted to *Materials Chemistry and Physics* (under revision following the referee suggestions).

4. Chapter 5: "Theoretical studies on interactions of glucose with hydrated ZnO surfaces", **Sasthi Charan Mandal**, Jaydeb Chakrabarti, *Manuscript in preparation*.

1.2 References

(1) Dowd, D. R.; Lloyd, R. S. Biological Significance of Facilitated Diffusion in Protein-DNA Interactions. Applications to T4 Endonuclease V-Initiated DNA Repair. *J. Biol. Chem.* **1990**, *265* (6), 3424–3431.

(2) González-Ruiz, V.; Olives, A. I.; Martín, M. A.; Ribelles, P.; Ramos, M. T.; Menéndez, J. C. An Overview of Analytical Techniques Employed to Evidence Drug-DNA Interactions. Applications to the Design of Genosensors. *Biomed. Eng. Trends Res. Technol.* **2011**, *32*, 215–219.

(3) Muthuchamy, M.; Muneeswaran, T.; Rajivgandhi, G.; Franck, Q.; Muthusamy, A.; Ji-Ming, S. Biologically Synthesized Copper and Zinc Oxide Nanoparticles for Important Biomolecules Detection and Antimicrobial Applications. *Mater. Today Commun.* **2020**, *22*, 100766.

(4) Saleh, T. A.; Fadillah, G. Recent Trends in the Design of Chemical Sensors Based on Graphene–Metal Oxide Nanocomposites for the Analysis of Toxic Species and Biomolecules. *TrAC Trends Anal. Chem.* **2019**, *120*, 115660.

(5) Solanki, P. R.; Kaushik, A.; Agrawal, V. V.; Malhotra, B. D. Nanostructured Metal Oxide-Based Biosensors. *NPG Asia Mater.* **2011**, *3* (1), 17–24.

(6) Pabo, C. O.; Sauer, R. T. Protein-DNA Recognition. *Annu. Rev. Biochem.* **1984**, *53* (1), 293–321.

(7) Pandey, P.; Hasnain, S.; Ahmad, S. Protein-DNA Interactions. **2019**.

(8) Phogat, N.; Kohl, M.; Uddin, I.; Jahan, A. Interaction of Nanoparticles with Biomolecules, Protein, Enzymes, and Its Applications. In *Precision Medicine*; Elsevier, 2018; pp 253–276.

(9) Sezer, A. D. *Application of Nanotechnology in Drug Delivery*; BoD–Books on Demand, 2014.

(10) Limo, M. J.; Sola-Rabada, A.; Boix, E.; Thota, V.; Westcott, Z. C.; Puddu, V.; Perry, C. C. Interactions between Metal Oxides and Biomolecules: From Fundamental Understanding to Applications. *Chem. Rev.* **2018**, *118* (22), 11118–11193.

(11) Pingoud, A.; Alves, J.; Geiger, R. Restriction Enzymes. In *Enzymes of Molecular Biology*; Springer, 1993; pp 107–200.

(12) Aggarwal, A. K. Structure and Function of Restriction Endonucleases. *Curr. Opin. Struct. Biol.* **1995**, *5* (1), 11–19.

(13) Loenen, W. A.; Dryden, D. T.; Raleigh, E. A.; Wilson, G. G.; Murray, N. E. Highlights of the DNA Cutters: A Short History of the Restriction Enzymes. *Nucleic Acids Res.* **2013**, *42* (1), 3–19.

(14) Smith, H. O.; Nathans, D. A Suggested Nomenclature for Bacterial Host Modification and Restriction Systems and Their Enzymes. *J. Mol. Biol.* **1973**, *81* (3), 419–423.

(15) Lynch, I.; Dawson, K. A. Protein-Nanoparticle Interactions. *Nano Today* **2008**, *3* (1–2), 40–47.

(16) Huang, X.; El-Sayed, I. H.; Qian, W.; El-Sayed, M. A. Cancer Cell Imaging and Photothermal Therapy in the Near-Infrared Region by Using Gold Nanorods. *J. Am. Chem. Soc.* **2006**, *128* (6), 2115–2120.

(17) Huang, X.; El-Sayed, I. H.; Qian, W.; El-Sayed, M. A. Cancer Cells Assemble and Align Gold Nanorods Conjugated to Antibodies to Produce Highly Enhanced, Sharp, and Polarized Surface Raman Spectra: A Potential Cancer Diagnostic Marker. *Nano Lett.* **2007**, *7* (6), 1591–1597.

(18) Whaley, S. R.; English, D. S.; Hu, E. L.; Barbara, P. F.; Belcher, A. M. Selection of Peptides with Semiconductor Binding Specificity for Directed Nanocrystal Assembly. *Nature* **2000**, *405* (6787), 665.

(19) Nel, A.; Xia, T.; Mädler, L.; Li, N. Toxic Potential of Materials at the Nanolevel. *science* **2006**, *311* (5761), 622–627.

(20) Vertegel, A. A.; Siegel, R. W.; Dordick, J. S. Silica Nanoparticle Size Influences the Structure and Enzymatic Activity of Adsorbed Lysozyme. *Langmuir* **2004**, *20* (16), 6800–6807.

(21) Pingoud, A.; Fuxreiter, M.; Pingoud, V.; Wende, W. Type II Restriction Endonucleases: Structure and Mechanism. *Cell. Mol. Life Sci.* **2005**, *62* (6), 685.

(22) Pingoud, A.; Wilson, G. G.; Wende, W. Type II Restriction Endonucleases—a Historical Perspective and More. *Nucleic Acids Res.* **2014**, *42* (12), 7489–7527.

(23) Baldwin, G. S.; Sessions, R. B.; Erskine, S. G.; Halford, S. E. DNA Cleavage by the EcoRV Restriction Endonuclease: Roles of Divalent Metal Ions in Specificity and Catalysis. *J. Mol. Biol.* **1999**, *288* (1), 87–103.

(24) Jeltsch, A.; Alves, J.; Wolfes, H.; Maass, G.; Pingoud, A. Substrate-Assisted Catalysis in the Cleavage of DNA by the EcoRI and EcoRV Restriction Enzymes. *Proc. Natl. Acad. Sci.* **1993**, *90* (18), 8499–8503.

(25) Kostrewa, D.; Winkler, F. K. Mg2+ Binding to the Active Site of EcoRV Endonuclease: A Crystallographic Study of Complexes with Substrate and Product DNA at 2-. ANG. Resolution. *Biochemistry* **1995**, *34* (2), 683–696.

(26) Robinson, C. R.; Sligar, S. G. Changes in Solvation during DNA Binding and Cleavage Are Critical to Altered Specificity of the EcoRI Endonuclease. *Proc. Natl. Acad. Sci.* **1998**, *95* (5), 2186–2191.

(27) Jeltsch, A.; Alves, J.; Maass, G.; Pingoud, A. On the Catalytic Mechanism of EcoRI and EcoRV A Detailed Proposal Based on Biochemical Results, Structural Data and Molecular Modelling. *FEBS Lett.* **1992**, *304* (1), 4–8.

(28) Robinson, C. R.; Sligar, S. G. *Molecular Recognition Mediated by Bound Water: A Mechanism for Star Activity of the Restriction Endonuclease EcoRI*; Elsevier, 1993.

(29) Alim, K. A.; Fonoberov, V. A.; Shamsa, M.; Balandin, A. A. Micro-Raman Investigation of Optical Phonons in ZnO Nanocrystals. *J. Appl. Phys.* **2005**, *97* (12), 124313.

(30) Saniz, R.; Sarmadian, N.; Partoens, B.; Batuk, M.; Hadermann, J.; Marikutsa, A.; Rumyantseva, M.; Gaskov, A.; Lamoen, D. First-Principles Study of CO and OH Adsorption on In-Doped ZnO Surfaces. *J. Phys. Chem. Solids* **2019**, *132*, 172–181.

(31) Chandraboss, V. L.; Karthikeyan, B.; Senthilvelan, S. Experimental and First-Principles Study of Guanine Adsorption on ZnO Clusters. *Phys. Chem. Chem. Phys.* **2014**, *16* (42), 23461–23475.

(32) Gong, M.; Yang, Z.; Xu, X.; Jasion, D.; Mou, S.; Zhang, H.; Long, Y.; Ren, S. Superhydrophobicity of Hierarchical ZnO Nanowire Coatings. *J. Mater. Chem. A* **2014**, *2* (17), 6180–6184.

(33) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis. *Annu. Rev. Biochem.* **2011**, *80*, 669–702.

(34) Kamerlin, S. C.; Sharma, P. K.; Prasad, R. B.; Warshel, A. Why Nature Really Chose Phosphate. *Q. Rev. Biophys.* **2013**, *46* (1), 1–132.

(35) Lodish, H.; Zipursky, S. L. Molecular Cell Biology. *Biochem Mol Biol Educ* **2001**, *29*, 126–133.

(36) Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P.; Wilson, J.; Hunt, T. *Molecular Biology of the Cell*; WW Norton & Company, 2017.

(37) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Molecular Biology of the Cell. **2007**.

(38) Kutzelnigg, W.; von Herigonte, P. Electron Correlation at the Dawn of the 21st Century. In *Advances in quantum chemistry*; Elsevier, 2000; Vol. 36, pp 185–229.

(39) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785.

(40) Ventura, E. E.; Davis, J. N.; Goran, M. I. Sugar Content of Popular Sweetened Beverages Based on Objective Laboratory Analysis: Focus on Fructose Content. *Obesity* **2011**, *19* (4), 868–874.

(41) Chen, C.; Xie, Q.; Yang, D.; Xiao, H.; Fu, Y.; Tan, Y.; Yao, S. Recent Advances in Electrochemical Glucose Biosensors: A Review. *Rsc Adv.* **2013**, *3* (14), 4473–4491.

CHAPTER 2

Specificity of Metal Ion Cofactor in DNA Cleavage by EcoRV

2.1 Introduction

Restriction endonucleases (RE) catalyse the cleavage of DNA. These enzymes are not only important to protect bacterial cells from invasion by viral DNA,[1–3] but also widely used as genetic tools.[4,5] The REs require metal ions as cofactor for their activities, Mg^{2+} being the natural cofactors for majority of enzymes, like EcoAI [6], EcoRI [7], EcoRV [7–10], SepMI [10,11], EhoI [10,11] to name only a few. Although REs can bind to other divalent metal ions, the protein has specific binding [11] to DNA in presence of both Ca^{2+} and Mg^{2+}. However, gel shift experiments show that the REs do not show DNA cleavage in presence of Ca^{2+} in contrast to Mg^{2+}.[12,13] Numerous studies have been performed to understand role of Mg^{2+} as natural cofactor. [11,14–17] But the microscopic origin of the specificity of metal ion cofactor Mg^{2+} *vis-à-vis* Ca^{2+} in DNA cleavage is unknown.[18–20]

Naturally occurring restriction endonucleases are classified as four types such as type I, II, III and IV, depending on their structure, recognition site, cleavage site etc.[1,2] Over 3000 type II restriction endonucleases have been identified, having more than 200 different DNA sequence specificities.[8] Type II REs recognize 4-8 base pair long DNA sequence.[8] The best characterized RE is EcoRV, a type II restriction endonuclease from Escherichia coli.[21] EcoRV specifically binds to recognition sequence GATATC in DNA in presence of Mg^{2+} and Ca^{2+}. It cleaves at the central base pair step (TA-AT) only in presence of Mg^{2+}.[22] Crystal structures of EcoRV and DNA complexes with two protein subunits in presence of both Mg^{2+} [23] and Ca^{2+} [24] have been reported. These structures reveal that the cognate DNA is in bent conformation in both cases. However, there is difference in the two cases so far the location of the metal ions

are concerned. In case of Mg^{2+} ion bound complex, both the metal ions are present in the same subunit, where one of the ions directly interacts with DNA backbone. In case of Ca^{2+}, the ions are symmetrically located on both the subunits, enabling both to directly interact with the DNA backbone. The acidic amino acids aspartic acid, ASP74 and ASP90 holding the metal ion and basic amino acid lysine, LYS92 of EcoRV lie in the vicinity of cleavage site. These residues are important in cleavage process as shown by several experimental [25–28] and earlier theoretical studies.[29] Despite similarity in binding residue type, DNA cleaving takes place in presence of Mg^{2+} and not in presence of Ca^{2+}. A recent experimental study on EcoRV-DNA complex shows that the speed of the enzymatic reaction can be modulated by mechanically stressing the DNA substrate.[30] The bent DNA conformations in the crystal structures in presence of both Mg^{2+} and Ca^{2+}, rule out that the overall DNA conformations account for the specificity of cofactor Mg^{2+} *vis-à-vis* Ca^{2+}.[18–20] This, however, does not rule out the possibility of sensitivity of fluctuations of microscopic conformational variables[31] to the type and locations of the metal ions. The microscopic conformational variables are the dihedral angles in protein and the relative displacements and orientations of the base pair steps in DNA. Our primary objective here is to explore if their fluctuations can yield physical insight to cofactor sensitivity of the REs.

Fluctuations of microscopic conformational variables of a macromolecule can be captured from in-silico molecular simulations.[32] If the simulated trajectory is sufficiently long, the distribution of fluctuations of a conformational degree of freedom can be taken as a marginal probability distribution of the given degree of freedom in the background of all other fluctuating conformational variables, accounting for all sorts of coupling with the other variables implicitly. If the marginal distributions of the microscopic conformational variables between two conformational states of the macromolecule are known in a given ensemble, one can compute the changes in conformational free energy using the Boltzmann weight and

entropy from the Gibb's formula, comparing the marginal distributions in two states [33,34]. This constitutes a mean field description[35] of thermodynamics of macromolecular conformation. If the correlations between the microscopic conformational variables of the binding partners in a bio-molecular complex are small, one can take the conformational contributions to the free energy and entropy simply additive and can assign the conformational contribution to free energy and entropy changes to each of the binding partners. In practice, in such cases, the complex and the individual binding partners can be simulated in separate simulation boxes in a given ensemble to obtain the conformational contribution to thermodynamics. Such estimates[33,34] agree to the conformational contribution to entropy in protein complexes, extracted from NMR experiments[36] and give useful information on stability of protein complexes[37] and protein functions[38–41]. Here we extend the thermodynamic analysis based on conformational variables to DNA-protein complex.

We denote the complexes as follows: Mg^{2+}(A)-EcoRV-DNA denoting complex with asymmetrically located Mg^{2+} ions, and Ca^{2+}(S)-EcoRV-DNA that with symmetrically located Ca^{2+} ions as in the crystal structures. We compute conformational thermodynamics data of the cognate DNA base pairs and protein residues in these complexes using long all-atom molecular dynamics (MD) simulation trajectories. We simulate the complex, individual DNA and metal ion bound protein in separate simulation boxes with same number (N) of particles, same pressure (P=1atm) and temperature (T=300 K). The particle number is kept fixed by adjusting the number of water molecules. We focus on the cleavage region. We observe that both Mg^{2+}(A)-EcoRV-DNA and Ca^{2+}(S)-EcoRV-DNA are conformationally stable and ordered compared to the free components. However, the DNA base pair steps and the protein residues show enhanced fluctuations in Ca^{2+}(S)-EcoRV-DNA complex compared to Mg^{2+}(A)-EcoRV-DNA complex so that the metal ion cannot be held in the proximity of the cleavage point in Ca^{2+}(S)-EcoRV-DNA in contrast to that in Mg^{2+}(A)-EcoRV-DNA. We also study model

systems, namely, Mg^{2+}(S)-EcoRV-DNA by replacing Ca^{2+} by Mg^{2+} and Ca^{2+}(A)-EcoRV-DNA by replacing Mg^{2+} by Ca^{2+}. Such replacement of the metal ions leads to conformational destabilization that hinder to stabilize the metal ion in proximity to the phosphate group at the cleavage point. Such differences in conformational fluctuations lead to differences of Mg^{2+} and Ca^{2+} in the role of cofactor for catalytic activity of EcoRV on substrate DNA.

2.2 Materials and methods

2.2.1 System preparations

The initial structures for all-atom MD simulations (See Methods for details) are taken from PDB files with cognate DNA bound to EcoRV in presence of Mg^{2+} (Mg^{2+}(A)-EcoRV-DNA, pdb id: 1RVB)[23] and Ca^{2+} (Ca^{2+}(S)-EcoRV-DNA, pdb id: 1B94)[24]. The sequence of DNA bases in the crystal structure is 5'-dAAAGATATCTT-3' in 5'-3' direction and the complementary sequence as per the Watson-Crick base pairs [42] in 3'-5' direction. EcoRV in both pdbs is a dimer of two identical chains, A and B, each chain having 244 residues. The crystal structure shows Mg^{2+} ions in the vicinity of cleavage point P_{18} in 3'-5' direction. A sharp kink with a roll angle of 50^0 has been found at the centre (TA) base pair when the cognate DNA is bound to EcoRV, while that of normal B-DNA is about 1.5^0.[43] The initial structure of free DNA, containing the recognition sequence (GATATC), is made using nucgen software.[44] The crystal structure of Mg^{2+} bound EcoRV, Mg^{2+}-EcoRV is not available. We take Mg^{2+} coordinates from the crystal structure of Mg^{2+}(A)-EcoRV-DNA complex and copy these coordinates into the crystal structure of free EcoRV (pdb id 1rve [22]) and take the resulting structure as the initial structure. We make Ca^{2+} bound EcoRV in the same way as we made Mg^{2+} bound EcoRV, since the crystal structure of Ca^{2+}-EcoRV is not available. We have also simulated model systems where we replace Mg^{2+} ions by Ca^{2+} ions in Mg^{2+}(A)-EcoRV-DNA complex, labelled as

Ca^{2+}(A)-EcoRV-DNA complex. Similarly, we replace Ca^{2+} ions by Mg^{2+} ions in Ca^{2+}(S)-EcoRV-DNA complex to form Mg^{2+}(S)-EcoRV-DNA.

2.2.2 Molecular dynamics (MD) simulations

We use Amberff14SB[45] force field for protein and the Amberff99bsc0[46] force field for DNA. The details of the method is described in Appendix I of this chapter. We prepare the systems using the standard protocol of the AmberTools 15[47] package. The systems are solvated in a cubic water box, containing TIP3P water molecules in such a way that there are at least 15 Å thick layers of water molecules around the solute. The periodic boundary conditions are imposed in all directions. By adding the required number of sodium (Na^+) and chloride (Cl^-) ions, the system is neutralized. Subsequently then the system is energy minimized using both steepest descent[48] and conjugate gradient[49] algorithms. Particle mesh Ewald summation method is used for long ranged electrostatic interaction with 1 Å grid spacing and 10^{-6} convergence criterion. The Lennard-Jones and the short-ranged electrostatic interactions are truncated at 10 Å.

Starting with the energy minimized structure, MD simulations are carried out using NAMD program[50] at 310 K and 1 atmospheric pressure in an isothermal–isobaric (NPT) ensemble. The temperature is maintained by using Langevin thermostat, and the pressure is controlled with a Nosé-Hoover Langevin piston. The SHAKE constraints are applied to all bonds involving hydrogen atoms. All calculations are done using a simulation time step of 1 fs.

2.2.3 DNA base pair step parameters

There are six DNA base pair step parameters, where three are rotational parameters (tilt (τ), roll (ρ), and twist(ω)) and other three are translational parameters (shift (Dx), slide (Dy), and rise (Dz)). Shift is defined as the displacement about an axis in the base-pair plane, directed

from minor to major groove, while slide is the displacement about an axis in base pair plane, directed from one strand to the other. Rise is described as the displacement along the helix axis. Tilt, roll, and twist are defined as the rotation of the base pairs around shift, slide, and rise axes. The DNA base pair step parameters are calculated using the formulae:

$\tau = - sin^{-1}(\bar{Z}_m . \bar{X}_1)$,

$\rho = sin^{-1}(\bar{Z}_m . \bar{Y}_1)$,

$\omega = cos^{-1}[(\bar{X}_1 \times \bar{Z}_m) . (\bar{X}_2 \times \bar{Z}_m)]$,

$Dx = \bar{M} . (\bar{X}_1 + \bar{X}_2)/|\bar{X}_1 + \bar{X}_2|$,

$Dy = \bar{M} . (\bar{Y}_1 + \bar{Y}_2)/|\bar{Y}_1 + \bar{Y}_2|$, and

$Dz = \bar{M} . \bar{Z}_m$

The coordinate frame consists of base pair (BP) short axis ($\bar{X}$) pointing towards the major groove, BP long axis ($\bar{Y}$) directed along the line joining C8 (of purine) and C6 (of pyrimidine) in the BP, and the mean of the normal of the bases in the BP ($\bar{Z}$). Here 1 and 2 denote two successive base pair planes. $\bar{M}$ is vector joining the BP centers of two consecutive base pairs. $\bar{Z}_m$ is defined in the following formula,

$$\bar{Z}_m = \frac{(\bar{X}_1 + \bar{X}_2) \times (\bar{Y}_1 + \bar{Y}_2)}{|(\bar{X}_1 + \bar{X}_2)| . |(\bar{Y}_1 + \bar{Y}_2)|} .$$

2.2.4 Conformational thermodynamics

The detailed description of histogram based method (HBM) for calculating the conformational thermodynamics is reported.[34] The normalized probability distribution of any conformational variable θ in free states and complex states are given by $H_i^{free}(\theta)$ and

$H_i^{complex}(\theta)$ respectively. The change in free energy of any dihedral θ of the bound state as compared to free state is defined as,

$$\Delta G_i^{conf}(\theta) = - k_B T \ln (H_{max,i}^{complex}(\theta) / H_{max,i}^{free}(\theta)),$$

where 'max' represents the peak value of the histogram and i represents the protein residue. The change in conformational entropy of a given dihedral θ, is evaluated from Gibbs formula,

$$\Delta S_i^{conf}(\theta) = - k_B [\sum_j H_{i,j}^{complex}(\theta) \ln H_{i,j}^{complex}(\theta) - \sum_j H_{i,j}^{free}(\theta) \ln H_{i,j}^{free}(\theta)].$$

Here the sum is taken over all histogram bins j. θ corresponds to the base pair step parameter for DNA and dihedral angle for the protein. The dihedral distributions are calculated using in-house program.

2.2.5 Identification of interfacial interactions

We calculate interactions between EcoRV and DNA atoms in the protein-DNA interface. The interface is defined as the region where a pair of atoms, one belonging to the DNA and the other to the protein, lie within 6 Å distance. This computation is done on the average structure of the molecules generated over the equilibrium portion of the trajectory.

The interactions are characterized by distance and angle criteria.[51] The hydrogen bonds are taken when the distance between donor (D) and acceptor (A) is less or equal to 3.4 Å and the angle (D-H-A) cut off is 160^0. For salt bridge (SB) interactions, the distance between any of the nitrogen atoms of basic residues and any of the oxygen atoms of phosphate group are within 3.4 Å and the cut off angle is same as in HB interaction. Ionic species belonging to the DNA and the protein at the interface are considered for electrostatic interactions. The calculations of hydrogen bond, electrostatics and salt bridge interactions are done using Discovery studio software.[51]

2.3 Results and discussion

The force field parameters for Mg^{2+} and Ca^{2+} are shown in Table 1. The total number of water molecules (see Table 2) is adjusted to ensure that the total number of atoms is the same in all cases in order to make the ensembles comparable.

Table 1. Lennard-Jones parameters for Mg^{2+} and Ca^{2+} ions.

Metal ion	R_{min} (Å) (= $2^{1/6}$ σ)	ε (kcal mol^{-1})
Mg^{2+}	0.7926	0.8947
Ca^{2+}	1.7131	0.4598

Table 2. The number of water molecules taken to solvate the cubical boxes and the total number of atoms in the systems.

Systems	No. of water molecules in the box	Total number of atoms
Mg^{2+}(A)-EcoRV-DNA	52164	165374
Ca^{2+}(S)-EcoRV-DNA	52164	165374
Mg^{2+}-EcoRV	52404	165374
Ca^{2+}-EcoRV	52404	165374
Ca^{2+}(A)-EcoRV-DNA	52164	165374
Mg^{2+}(S)-EcoRV-DNA	52164	165374
Free DNA	54885	165375

The numbering of phosphate, sugar and bases is shown in Fig. 1. The DNA contains the cognate sequence, consisting of G_4:C_{20}, A_5:T_{19}, T_6:A_{18}, A_7:T_{17}, T_8:A_{16} and C_9:G_{15}, shown in the figure by the box in dotted line. The cleavage occurs at P_7 and P_{18} between T_6:A_{18} and A_7:T_{17} base pairs, represented by arrows in both strands. We have carried out total simulation for 1μs for the systems, Mg^{2+}(A)-EcoRV-DNA, Ca^{2+}(S)-EcoRV-DNA, Mg^{2+}(A)-EcoRV, Ca^{2+}(S)-EcoRV and free DNA respectively. For the other two model systems, Mg^{2+}(S)-EcoRV-DNA and Ca^{2+}(A)-EcoRV-DNA, shorter simulations (250 ns) are carried out. The root mean square deviations for free DNA (Fig. 2(a)), Mg^{2+}-EcoRV (Fig. 2(b)), Ca^{2+}-EcoRV (Fig.

2(c)), Mg^{2+}(A)-EcoRV-DNA (Fig. 2(d)), Ca^{2+}(S)-EcoRV-DNA (Fig. 2(d)), Ca^{2+}(A)-EcoRV-DNA (Fig. 2(e)) and Mg^{2+}(S)-EcoRV-DNA (Fig. 2(f)) are used to judge the equilibrations of different systems. Using an equilibrated trajectory of 600-1000 ns for crystal structured complexes and 150-250 ns for model complexes, we compute histograms of protein dihedrals and histograms of DNA base pair step parameters.

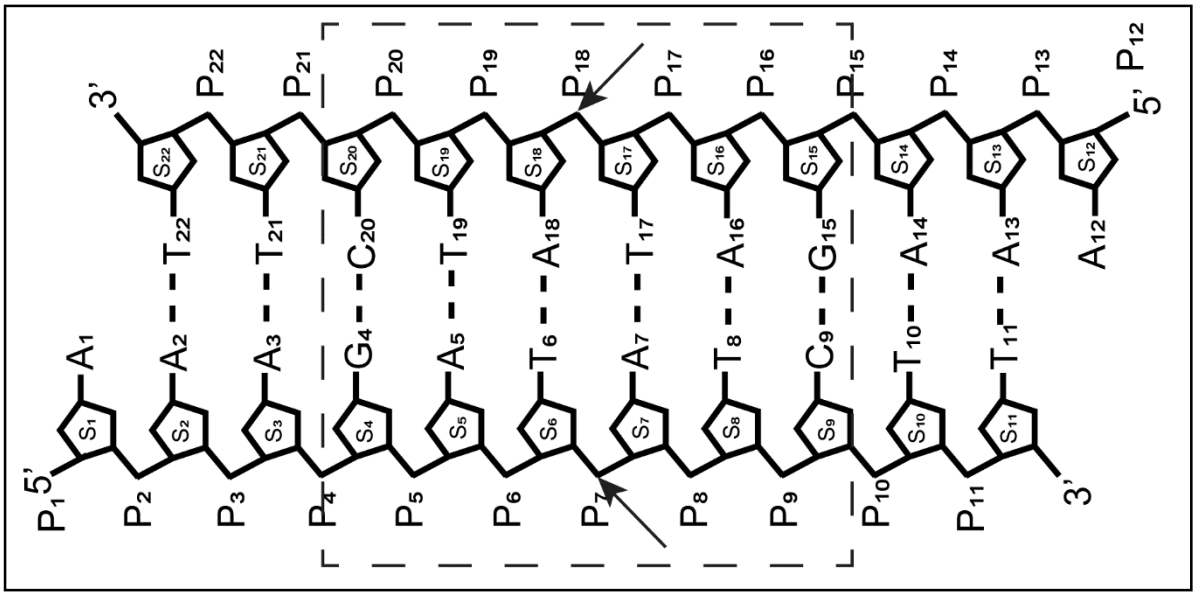

Fig. 1. Notations and numbering of DNA functional groups. The recognition sequence is marked by box of dashed line. The cleavage points are marked in arrows.

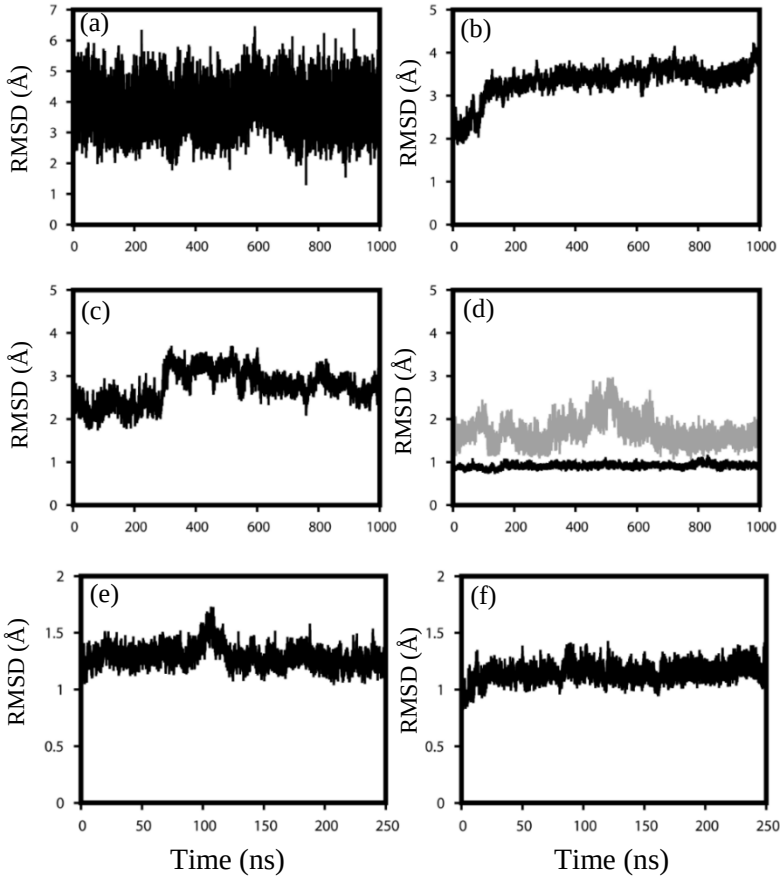

Fig. 2. Root mean square deviations (RMSD) (Å) of (a) Free DNA, (b) Mg^{2+}-EcoRV, (c) Ca^{2+}-EcoRV, (d) Mg^{2+}(A)-EcoRV-DNA (black), Ca^{2+}(S)-EcoRV-DNA (grey), (e) Ca^{2+}(A)-EcoRV-DNA and (f) Mg^{2+}(S)-EcoRV-DNA.

We first consider systems where the complexes are experimentally reported, namely, Mg^{2+}(A)-EcoRV-DNA and Ca^{2+}(S)-EcoRV-DNA and in addition, Mg^{2+}-EcoRV, Ca^{2+}-EcoRV, and the free DNA. We validate the simulated structures by comparing to the available crystal structures. The average structures for all the cases using equilibrated portion of the trajectories are generated. The DNA structure is given in terms of base pair step parameters (see Materials and methods for details), consisting of the rotational parameters, tilt (τ), roll (ρ), twist (ω) describing relative orientations of a base pair step defined by two successive planes, each consisting of a pair of bases and translational parameters D_x, D_y and D_z for relative displacements between the base pair planes in a given step [52]. The mean values of DNA base

pair step parameters of simulated structures for free DNA and DNA in complex are shown in Table 3. The values of the translational and rotational parameters are in agreement to the crystal structures.

Table 3. The values of DNA base pair step parameters corresponding to each base pair step of DNA in free and in Mg^{2+}(A)-EcoRV-DNA complex.

Base pairs	τ		ρ		ω		Dx		Dy		Dz	
	Free	Complex	Free	Complex	Free	Complex	Free	Complex	Free	Complex	Free	Complex
G_4:C_{20} A_5:T_{19}	-0.2	0.26	2.44	-14.57	34.01	35.66	0.06	-0.42	-0.8	-0.50	3.18	3.83
A_5:T_{19} T_6:A_{18}	-0.05	1.09	-0.46	1.237	32.17	22.77	-0.01	-0.36	-0.81	-0.55	3.4	3.99
T_6:A_{18} A_7:T_{17}	-0.06	1.89	11.25	55.57	30.5	24.80	-0.0	-0.35	-0.79	-0.59	3.2	3.6
A_7:T_{17} T_8:A_{16}	0.05	-1.94	-0.07	-8.04	31.93	31.74	0.08	-0.179	-0.82	-0.47	3.46	3.88
T_8:A_{16} C_9:G_{15}	0.96	1.05	2.0	-9.2	31.17	34.72	0.09	0.934	-0.56	-0.87	3.36	3.62
C_9:G_{15} T_{10}:A_{14}	-2.85	2.20	5.49	2.96	17.64	29.37	-0.96	-0.305	0.06	-1.58	3.5	3.22

We show the interface near the cleavage region as given in the Materials and methods section. The interface in Mg^{2+}(A)-EcoRV-DNA complex consists of T_6:A_{18} and A_7:T_{17} base pairs of DNA and ASP 36(A), LYS92(A) and ASN 188(A) residues from chain A and GLU 45(B), GLU 65(B), ASP74(B), ASP90(B) and LYS92(B) residues from chain B of EcoRV. Similarly, in Ca^{2+}(S)-EcoRV-DNA complex, the DNA base pairs belong to the interface are same as in Mg^{2+}(A)-EcoRV-DNA complex. The interfacial protein residues are ASP74(A), ASP90(A) and LYS92(A) from chain A and ASP74(B), ASP90(B) and LYS92(B) from chain B. The interfacial residues are in agreement to the crystal structures.

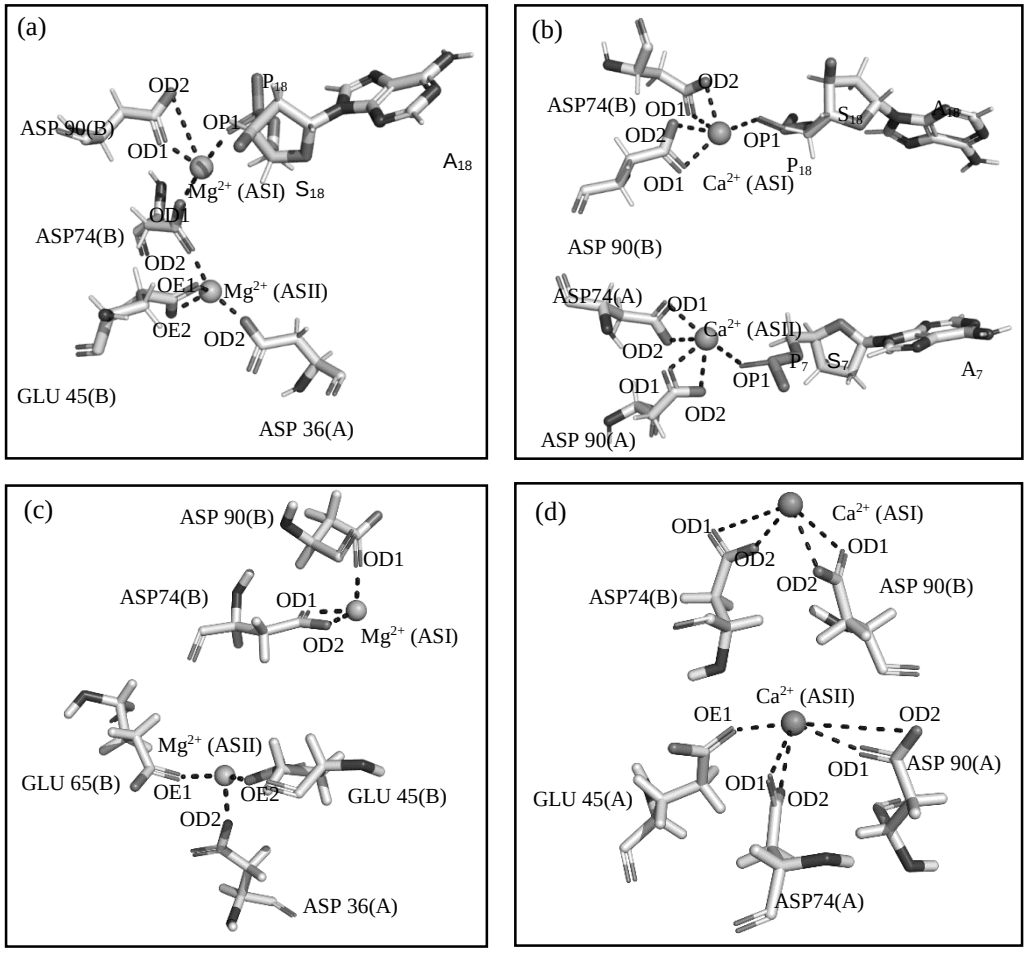

Fig. 3. (a) Coordinating residues of Mg^{2+}(ASI) and Mg^{2+}(ASII) in Mg^{2+}(A)-EcoRV-DNA complex and (b) coordinating residues of Ca^{2+} (ASI) and Ca^{2+} (ASII) in Ca^{2+}(S)-EcoRV-DNA from simulated structures. Coordinating residues of (c) Mg^{2+}(ASI) and Mg^{2+}(ASII) in Mg^{2+}(A)-EcoRV, and (d) Ca^{2+} (ASI) and Ca^{2+} (ASII) in Ca^{2+}(S)-EcoRV.

The Mg^{2+} coordinations in Mg^{2+}(A)-EcoRV-DNA complex are shown in Fig.3(a). There are two active sites for coordinating Mg^{2+}, denoted by Mg^{2+}(ASI) and Mg^{2+}(ASII). The coordination is considered if the distance between Mg^{2+} ions and the atoms of EcoRV residues or DNA is less than or equal to 3.0 Å. OD2 of ASP74(B), OD1 and OD2 of ASP90(B) and OP1 of P_{18} of DNA, participate in coordination to Mg^{2+}(ASI). For Mg^{2+}(ASII), OD2 of ASP 36(A), OE1 and OE2 of GLU 45(B), OD2 of ASP74(B) of EcoRV participate in coordination. In Ca^{2+}(S)-EcoRV-DNA the coordinating ligands (Fig. 3(b)) to Ca^{2+}(ASI) are: OD1 and OD2 of both ASP74(A) and ASP90(A), and OP1 of P_7. For Ca^{2+}(ASII), these are: OD1 and OD2 of

both ASP74(B) and ASP90(B), and OP1 of P_{18}. The coordinations in simulated structure are in agreement to the crystal structure.

In the simulated structure of Mg^{2+}-EcoRV, the coordinating residues to Mg^{2+}(ASI) are same as in the simulated structure of Mg^{2+}(A)-EcoRV-DNA complex, shown in Fig. 3(c). In case of Mg^{2+} (ASII), the coordinating residue ASP74(B) in Mg^{2+}(A)-EcoRV-DNA is replaced by GLU 65(B), while the rest of the coordinating residues are same (Fig. 3(c)). In Ca^{2+}-EcoRV, Ca^{2+} (ASI) is coordinated with OD1 and OD2 of both ASP74(B) and ASP90(B) (Fig. 3(d)). OD1 and OD2 of both ASP74(A) and ASP90(A), OE1 of GLU 45(A) coordinate to Ca^{2+} (ASII) (Fig. 3(d)). The rest of the coordinations in all the cases are fulfilled by water molecules.

2.3.1 Conformational stability and order of the complexes

We focus on the complexes having asymmetric locations of the metal ions in the crystal structures. The fluctuations in Mg^{2+}(A)-EcoRV-DNA complex are compared to those in free DNA and Mg^{2+}-EcoRV, while Ca^{2+}(S)-EcoRV-DNA complex to free DNA and Ca^{2+}-EcoRV in terms of microscopic variables, namely, dihedral angles in protein [31], and the base pair step parameters for DNA. In order to define these parameters, we consider each base and its Watson-Crick canonical pair to define a plane. The two consecutive base pairs (i, j) and (i+1, j+1) respectively, where i belongs to the first strand, and j belongs to the second strand of DNA, are defined as a single base pair step. Here, for the simplicity, we rewrite the above symbol as (i, i+1) to define a single base pair step. The details of the base pair step parameters are given in the Materials and methods. For a given base pair step parameter of base pair step (i, i+1) we denote the distribution by $H^{f}_{i,i+1}(b)$ in the free state and $H^{c}_{i,i+1}(b)$ in the complex c. Each peak in the histogram represents the most probable value of the base pair step parameters. Similarly, the histograms of dihedral θ of a protein residue R of the protein is given by $H^{(f)}_{R}(\theta)$ in free state and $H^{(c)}_{R}(\theta)$ in the complex c. The peaks in the backbone dihedral distributions denote

isomeric states, while those in the side chain dihedrals denote different rotameric states. Larger widths of the peaks denote more conformational fluctuations. We find that overall, the number of peaks, peak heights and widths are different in free and complexed states and typically sensitive to the nature of the metal ion.

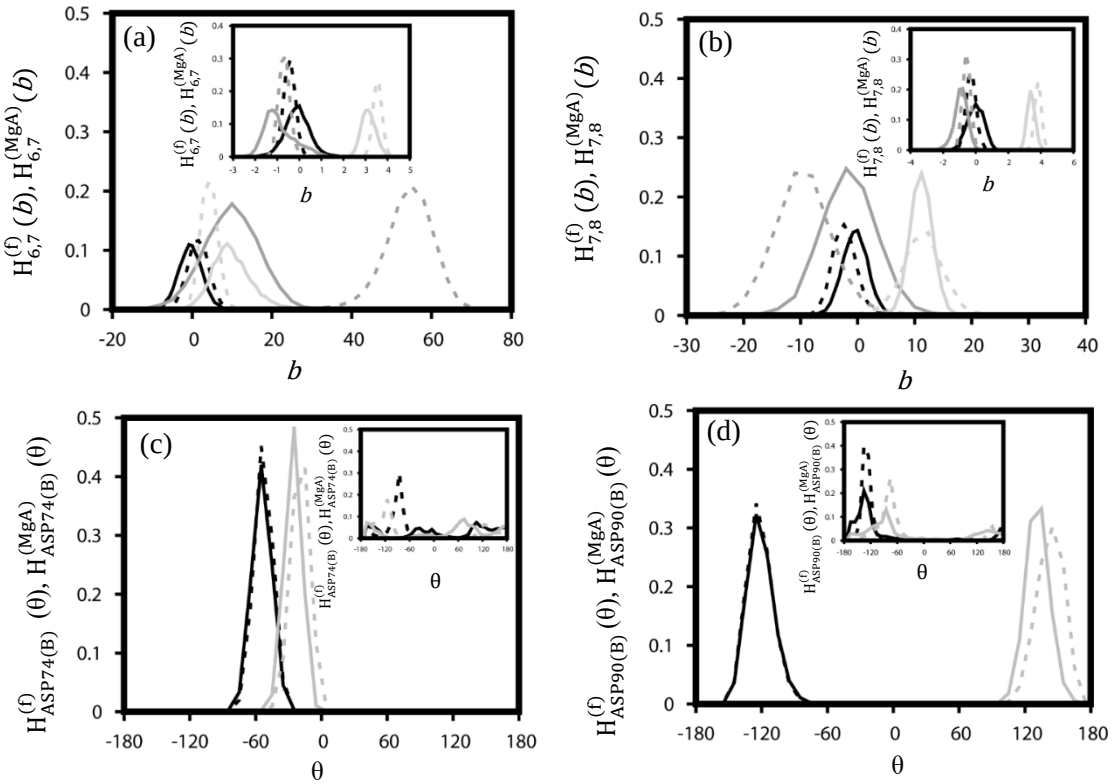

Fig. 4. The distributions $H^{(f)}_{i,i+1}(b)$ and $H^{(MgA)}_{i,i+1}(b)$ of base pair step parameters of base pair steps of DNA in free (solid lines) and in Mg^{2+}(A)-EcoRV-DNA (dotted lines). The histograms are: (a) τ (black), ρ (grey), ω (light grey) of (6,7). Inset: Histograms of the translations parameters: Dx(black), Dy(grey), Dz(light grey) of (6,7); (b) The histograms are: τ (black), ρ (grey), ω (light grey) of (7,8). Inset: Dx(black), Dy(grey), Dz (light grey) of (7,8). (c) The dihedral distributions of EcoRV in Mg^{2+}-EcoR, $H^{(f)}_R(\theta)$ (solid lines) and in Mg^{2+}(A)-EcoRV-DNA, $H^{(MgA)}_R(\theta)$ (dotted lines). The distributions are: φ(black), ψ(grey) of ASP74(B); and (d) φ(black), ψ(grey) of ASP90(B). The insets in (c) and (d) illustrate the distributions of χ_1(black), χ_2(grey) of ASP74(B) and ASP90(B).

We illustrate below a few cases of microscopic conformational degrees of freedom in the cleavage region of Mg^{2+}(A)-EcoRV-DNA complex. We first consider the fluctuations of the base pair step parameters of (6,7) and (7,8) base pair steps of DNA, and the fluctuations of dihedral angles of ASP74(B) and ASP90(B) residues of EcoRV. The distributions of the

rotational parameters for (6,7) base pair step are shown in Fig. 4(a). The distributions of τ for (6,7) base pair step are single peaked both in free and complex states. ρ shows single broad peaked distribution in the free state, while sharp single peaked distribution in the complex. The distributions of ω for both cases are single peaked. The inset in (A) shows the translational parameters of the same base pair steps. Both $H_{6,7}^{(f)}$(Dx) and $H_{6,7}^{(f)}$(Dy) have single broad peak, while $H_{6,7}^{(MgA)}$(Dx) and $H_{6,7}^{(MgA)}$(Dy) show sharp single peak. $H_{6,7}^{(f)}$(Dz) has broad single peak, while $H_{6,7}^{(MgA)}$(Dz) is having sharp single peak. For base pair step (7,8), the distributions for rotational parameters are shown in Fig. 4(b), while those for the translational parameters are in inset of Fig. 4(b). Here the distributions of τ and ρ have single broad peak for both free and complex states. The distributions of Dx, Dy, and Dz are all broader in free states than in complex. Few representative histograms for the protein are shown in Figs. 4(c) and (d). The backbone dihedral distributions of the residues, ASP74(B) and ASP90(B) in cleavage region are single peaked in both states (Fig. 4(c) and (d)). However, the side chains show differences. $H_{ASP74(B)}^{(f)}(\chi_1)$ and $H_{ASP74(B)}^{(f)}(\chi_2)$ are multimodal distributions, while $H_{ASP74(B)}^{(MgA)}(\chi_1)$ and $H_{ASP74(B)}^{(MgA)}(\chi_2)$ are single peaked, shown in inset of Fig. 4(c). χ_1 and χ_2 of ASP90(B) have multimodal distributions in free state and single peaked distributions in complex state, shown in inset of Fig. 4(d).

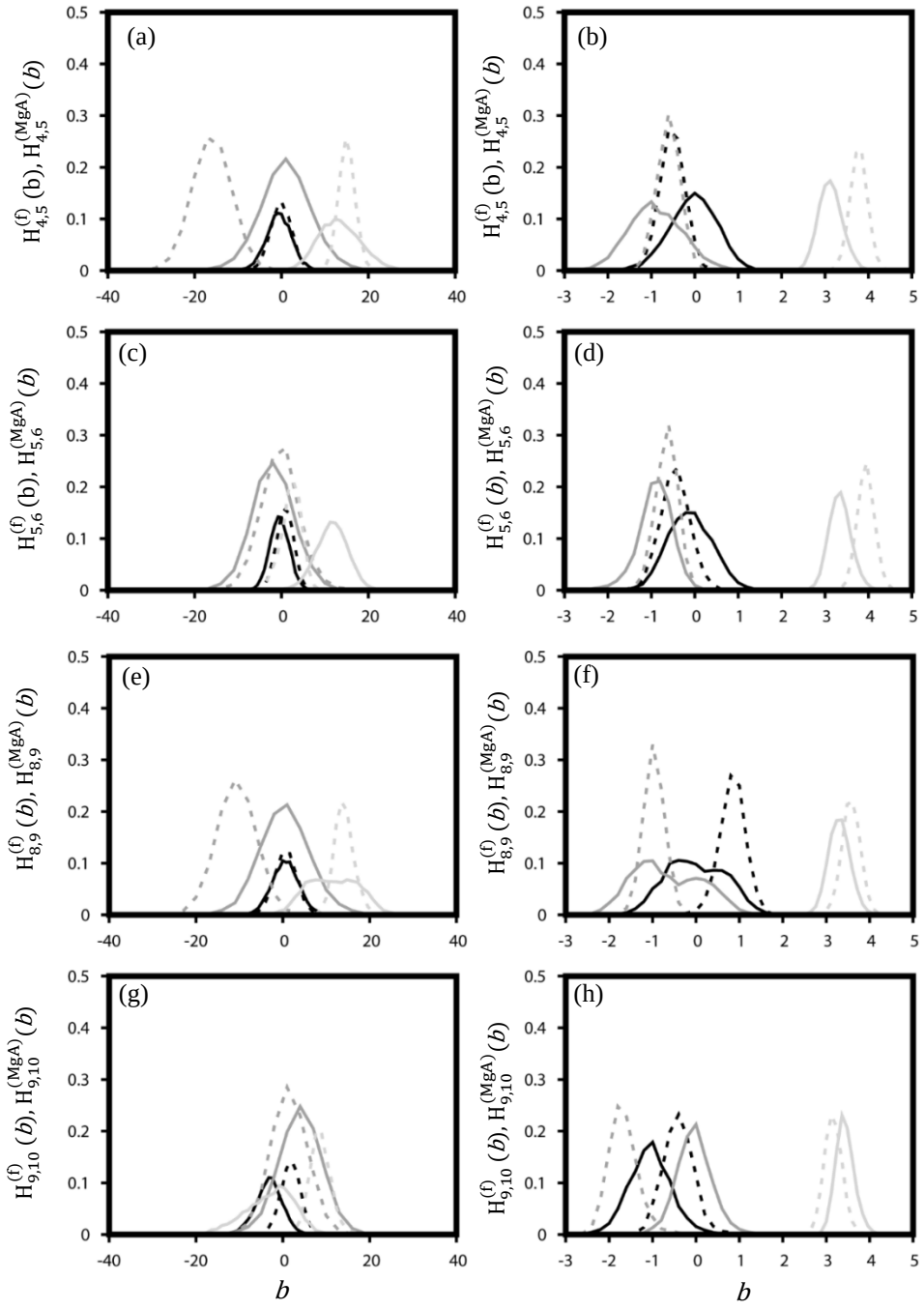

Fig. 5. Distributions of DNA base pair step parameters in Mg^{2+}(A)-EcoRV-DNA (solid lines) and in free DNA (dotted lines). (a) τ, ρ and ω and (b) Dx, Dy, and Dz of (4,5); (c) τ, ρ and ω and (d) Dx, Dy, and Dz of (5,6); (e) τ, ρ and ω and (f) Dx, Dy, and Dz of (8,9); (g) τ, ρ and ω and (h) Dx, Dy, and Dz of (9,10). The color conventions are: τ (black), ρ (grey), ω (light grey), Dx(black), Dy(grey), and Dz(light grey).

The histograms of the other base pair steps, (4,5), (5,6), (8,9) and (9,10) in recognition region of DNA in Mg^{2+}(A)-EcoRV-DNA complex along with the data for free DNA are shown in Figs. 5(a)-(h). All the rotational and translational parameters of (4,5) (Fig. 5(a) and (b)) and

(5,6) (Fig. 5(c) and (d)) base pair steps show single peaked distributions both in free state and Mg^{2+}(A)-EcoRV-DNA complex. τ and ρ of (8,9) are single peaked in both states, while ω has single peaked distribution in free state and multimodal distribution in the complex, shown in Fig. 5(e). Both Dx and Dy show single peaked distributions in free state and multimodal distributions in the complex, shown in Fig. 5(f). Dz is single peaked in both states (see Fig. 5(f)). All the base pair step parameters for (9,10) (see Fig. 5(g) and (h)) show single peaked distributions both in free and complex states.

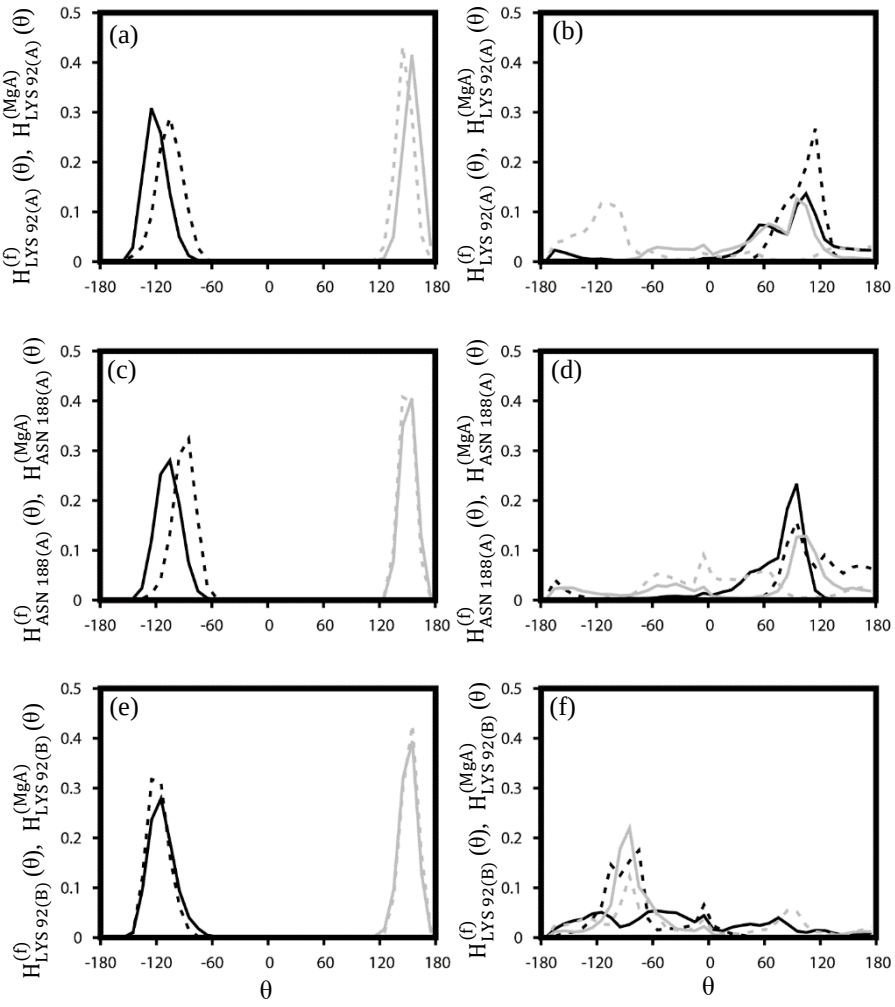

Fig. 6. The dihedral distributions of Mg^{2+}-EcoRV (solid lines) and Mg^{2+}(A)-EcoRV-DNA (dotted lines). (a) φ (black), ψ (grey) and (b) χ_1 (black), χ_2 (grey) of LYS 92(A); (c) φ (black), ψ (grey) and (d) χ_1 (black), χ_2 (grey) of ASN 188(A); (e) φ (black), ψ (grey) and (f) χ_1 (black), χ_2 (grey) of LYS 92(B).

The histograms of the other residues, LYS 92(A), ASN 188(A) and LYS 92(B) in cleavage region of Mg^{2+}(A)-EcoRV-DNA complex are shown in Figs. 6(a)-(f). We observe that in both the free and complex states, all the backbone dihedral angles of LYS 92(A) (Fig. 6(a)), ASN 188(A) (Fig. 6(c)), and LYS 92(B) (Fig. 6(e)) residues have single peaked distributions. Similarly, the side chain dihedral angles of LYS 92(A) (Fig. 6(b)), ASN 188(A) (Fig. 6(d)), and LYS 92(B) (Fig. 6(f)) residues show multimodal distributions both in free and complex.

The histograms for the DNA base pair step parameters and the protein dihedral angles in Ca^{2+}(S)-EcoRV-DNA (Figs. 7(a)-(d)) also show sensitivities to the complexed and free conformations. $H_{6,7}^{(f)}(\tau)$ is single peaked in free state, while $H_{6,7}^{(CaS)}(\tau)$ shows multimodal distribution for (6,7) base pair step (Fig. 7(a)). ρ and ω are broad single peaked in both free and complex states (Fig. 7(a)). $H_{6,7}^{(f)}(Dx)$ is single peaked distribution, while $H_{6,7}^{(CaS)}(Dx)$ is multimodal distribution, shown in inset of Fig. 7(a). Dy shows single peak with long tail in complex states, while sharp single peak in free state (see inset of Fig. 7(a)). Dz in both free and complex states show sharp single peak. For the base pair step (7,8), both τ and ω are sharp single peaked distributions, while ρ is broad single peak in both cases, shown in Fig. 7(b). The translational parameters, shown in inset of Figs. 7(b) are all single peaked in both cases. The histograms of the protein residues in cleavage region are shown in Figs. 7(c) and (d). The backbone dihedrals of ASP74(B) and ASP90(B) (Figs. 7(c) and (d)) are all single peaked. $H_{ASP74(B)}^{(f)}(\chi_1)$ and $H_{ASP74(B)}^{(f)}(\chi_2)$ are multimodal distributions, shown in inset of Fig. 7(c). Similarly, $H_{ASP74(B)}^{(CaS)}(\chi_1)$ and $H_{ASP74(B)}^{(CaS)}(\chi_2)$ are single peaked. The histograms of χ_1 and χ_2 in ASP90(B) (see inset of Fig. 7 (d)) show single peaked distributions in free state and multimodal distributions in complex state.

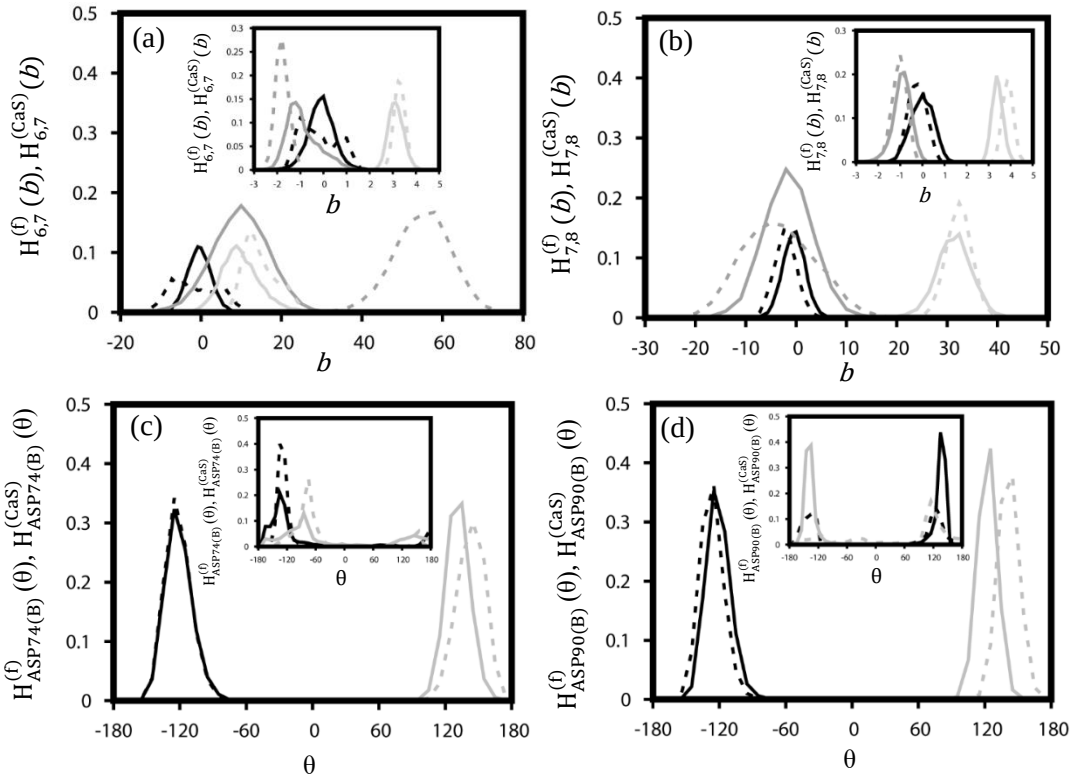

Fig. 7. The distributions $H^{(f)}_{i,i+1}(b)$ and $H^{(MgA)}_{i,i+1}(b)$ of base pair step parameters of base pair steps of DNA in free (solid lines) and in Ca^{2+}(S)-EcoRV-DNA (dotted lines). The histograms are: (a) τ, ρ, ω of (6,7). Inset: Histograms of the translational parameters: Dx, Dy, Dz of (6,7); (b) The histograms are: τ, ρ, ω of (7,8). Inset: Dx, Dy, Dz of (7,8). (c) The dihedral distributions of EcoRV in Mg^{2+}-EcoR, $H^{(f)}_{R}(\theta)$ (solid lines) and in Mg^{2+}(A)-EcoRV-DNA, $H^{(CaS)}_{R}(\theta)$ (dotted lines). The distributions are: φ, ψ of ASP74(B); and (D) φ, ψ of ASP90(B). The insets in (c) and (d) illustrate the distributions of χ_1, χ_2 of ASP74(B) and ASP90(B). The color conventions are: τ (black), ρ (grey), ω (light grey), Dx(black), Dy(grey) and Dz(light grey) for DNA base pair step parameters, and φ(black), ψ(grey), χ_1 (black) and χ_2 (grey) for protein dihedral angles.

The distributions of the backbone and side chain dihedral angles of the other residues, ASP74(A), ASP90(A), LYS92(A) and LYS92(B) in cleavage region of Ca^{2+}(S)-EcoRV-DNA complex are shown in Figs. 8(a)-(d). The backbone dihedral angles (φ and ψ) of ASP74(A) (Figs. 8(a)) show single peaked distributions both in free and complex states. χ_1 of ASP74(A) shows multimodal distributions both in free and complex, shown in inset of Fig. 8(a). $H^{(f)}_{ASP74(A)}(\chi_2)$ is single peaked distribution, while $H^{(CaS)}_{ASP74(A)}(\chi_2)$ shows multimodal distribution (see inset of Fig. 8(a)). φ and ψ of ASP90(A) (Figs. 8(b)) are all single peaked both in free and complex states. $H^{(f)}_{ASP90(A)}(\chi_1)$ and $H^{(f)}_{ASP90(A)}(\chi_2)$ are single peaked, while $H^{(CaS)}_{ASP90(A)}(\chi_1)$ and

$H^{(CaS)}_{ASP90(A)}$ (χ_2) are multimodal distributions, shown in inset of Fig. 8(b). The backbone dihedral angles of LYS92(A) (Figs. 8(c)) show single peaked both in free and complex states. $H^{(f)}_{LYS92(A)}$ (χ_1) is single peaked, while $H^{(CaS)}_{LYS92(A)}$ (χ_1) shows multimodal distribution (see inset of Fig. 8(c)). Both $H^{(f)}_{LYS92(A)}$ (χ_2) and $H^{(CaS)}_{LYS92(A)}$ (χ_2) are multimodal distributions (see inset of Fig. 8(c)). Both φ and ψ of LYS92(B) show single peaked distributions both in free and complex states, shown in Fig. 8(d). χ_1 and χ_2 of LYS92(B) show single peaked distributions in free state and multimodal distributions in the complex, shown in inset of Fig. 8(d).

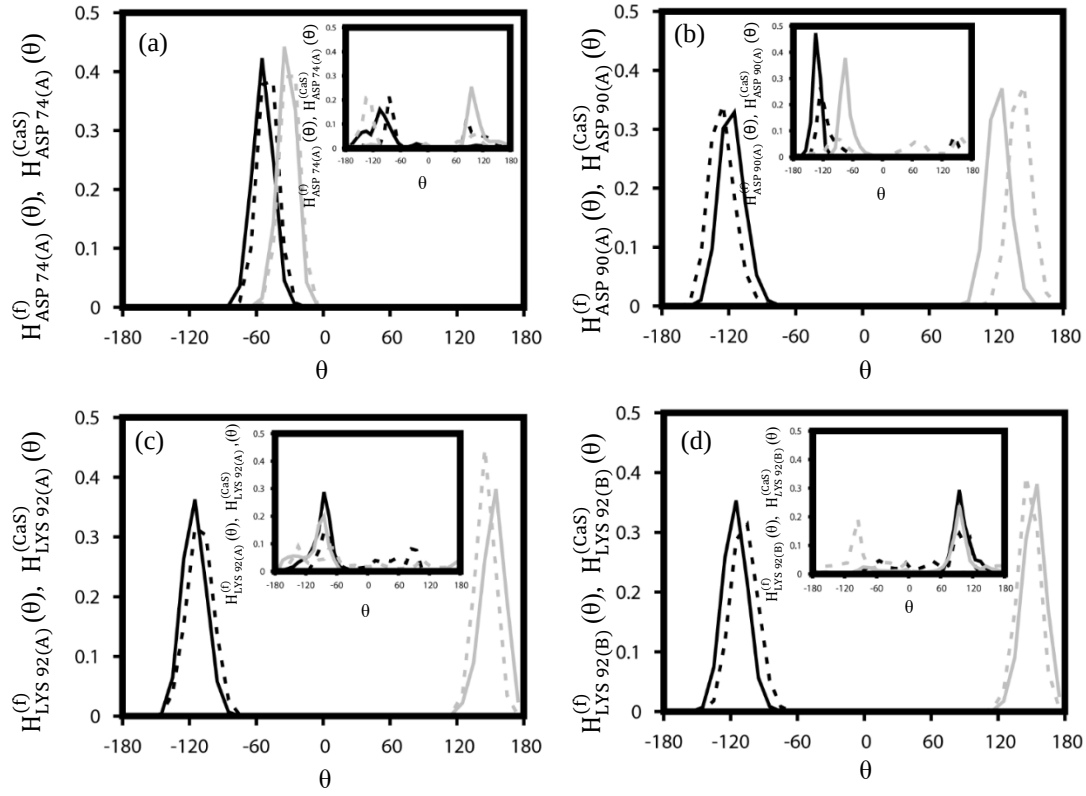

Fig. 8. The dihedral distributions for Ca^{2+}-EcoRV (solid lines) and for Ca^{2+}(S)-EcoRV-DNA (dotted lines). φ(black), ψ(grey) of (a) ASP 74(A), (b) ASP 90(A), (c) LYS 92(A) and (d) LYS 92(B). The insets illustrate the distributions of corresponding χ_1(black) and χ_2(grey).

In order to compute the thermodynamics due to conformational changes, we compute the statistical correlation[53,54] between all the variables where we find that the correlation between the base pair step parameters and the protein dihedral angles are small (correlation coefficient

< 0.5). Hence, we compute the changes in conformational free energy and entropy (see Materials and methods for details) in the complex with respect to free states of the components, simulated in equivalent ensemble. We consider the individual histograms as marginal distributions, implicitly taking care of coupling between the conformational degrees of freedom of the individual systems. We check the convergence of the marginal distributions by drawing 20 random samples, each consisting of 500 conformations over the equilibrated trajectories and compute the thermodynamic quantities. We find that the values are in agreement within an error of $\pm$5% in the mean. We arrive at similar estimate of error in the mean by constructing data for consecutive windows each of 50 ns width. We consider the mean values for further discussion. The changes in conformational free energy and entropy for a given base pair step parameter for (i, i+1) step with respect to the free state are denoted by $\Delta G^{fc}_{i,i+1}(b)$ and $T\Delta S^{fc}_{i,i+1}(b)$ respectively. They are conformational thermodynamic cost by the DNA base pairs for going from free to complexed state. The free energy and entropy changes of EcoRV residues in complex c with respect to the free state are denoted as $\Delta G_i^{conf}(\theta)$ and $T\Delta S_i^{conf}(\theta)$, where i denotes the EcoRV residue and θ, a dihedral angle of the residue. These quantities are conformational thermodynamic cost of the protein residues to go from free to complexed states. A negative change in free energy and entropy represents conformational stability and ordering respectively in the complex with respect to the free components, whereas positive changes represent conformational destabilization and disorder respectively.

The data for each base pair step in recognition sequence in the Mg^{2+}(A)-EcoRV-DNA complex with respect to the free DNA $\Delta G^{fMgA}_{i,i+1}(b)$ and $T\Delta S^{fMgA}_{i,i+1}(b)$ are shown in Fig. 9(a) and (b) respectively. The maximum conformational stabilization occurs at (4,5) step at the beginning of the cognate sequence. The free energy of conformational stabilization then decreases till (9,10) step at the end of the cognate sequence with oscillation at every alternate

base pair step, shown in Fig. 9(a). In case of (4,5), the free energy change gets large contributions from ω and Dy, while other base pair step parameters contribute almost comparably. All the base pair step parameters contribute almost equally for (5,6) step. The base pair step (6,7) at the cleavage site is stabilized through the translational base pair step parameters. In case of (7,8), all base pair step parameters contribute almost equally. For (8,9), large contributions of free energy come from Dx only, while other base pair step parameters have almost equal contributions. Similarly, for (9,10), the base pair step parameters contribute almost equally. Thus, the translational base pair step parameters contribute more than the rotational parameters, and make uneven contributions to different base steps, leading to oscillation in conformational free energy changes. We observe the oscillation in total change in conformational entropy as well. The conformational entropy decreases and is comparable at (4,5) and (6,7) (Fig. 9(b)), containing the cleavage site and (8,9), suggesting enhanced ordering in the complex than in the free DNA. The individual contributions of thermodynamic changes of dihedrals of the protein residues in the cleavage region and the metal ion coordinating residues $\Delta G_i^{conf}(\theta)$ and $T\Delta S_i^{conf}(\theta)$ are shown in Table 4. We find that the residues except ASP74(B), ASP90(B) and LYS92(B) show marginal changes (~ 2.5 kJ/mol) in conformational thermodynamics. ASP74(B) and ASP90(B) coordinate with Mg^{2+} ions are ordered through reduced flexibility of the side chain dihedrals. LYS92(B) in cleavage region is disordered through enhanced fluctuations in the side chain dihedrals.

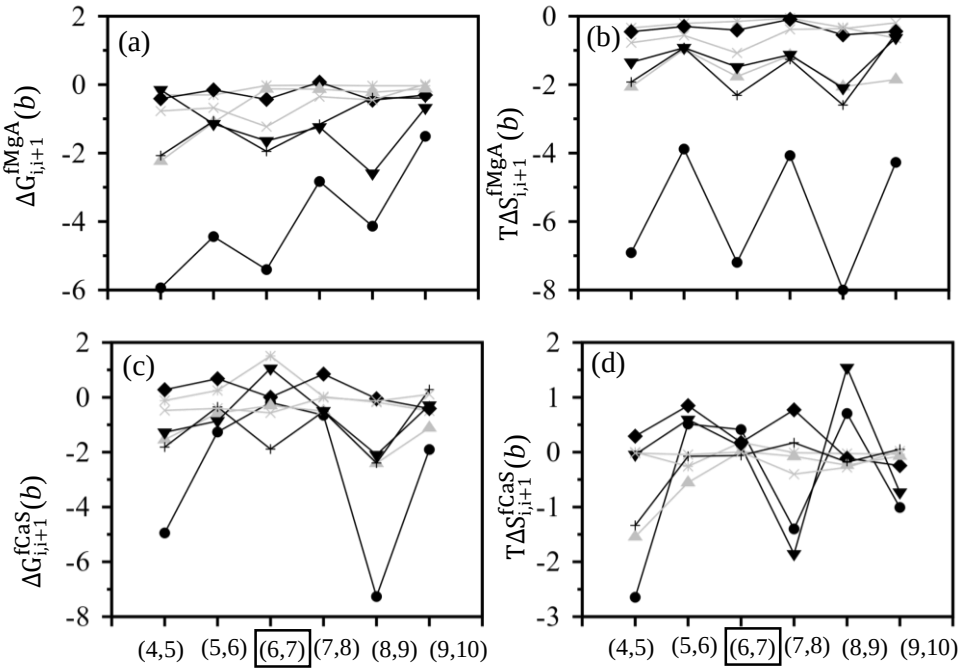

Fig. 9. (a) $\Delta G^{fMgA}_{i,i+1}(b)$ and (b) $T\Delta S^{fMgA}_{i,i+1}(b)$ of DNA base pair step parameters in recognition region for DNA in Mg^{2+}(A)-EcoRV-DNA with respect to free DNA. (c) $\Delta G^{fCaS}_{i,i+1}(b)$ and (d) $T\Delta S^{fCaS}_{i,i+1}(b)$ for Ca^{2+}(S)-EcoRV-DNA. The circles connected by solid black lines are total change in thermodynamic parameters of each base pair. The symbols for individual base pair step parameters: τ (square), ρ (diamond), ω (up triangle), Dx (down triangle), Dy (cross) and Dz (diagonal cross). The cleavage base pair has been marked in box. All quantities are in kJ/mol.

Table 4. Changes in conformational free energy and entropy (kJ/mol) of the residues in cleavage region in Mg^{2+}(A)-EcoRV-DNA with respect to Mg^{2+}-EcoRV.

Residues	ΔG_i^{conf} (T)	φ	ψ	χ_1	χ_2	χ_3	χ_4	$T\Delta S_i^{conf}$ (T)	φ	Ψ	χ_1	χ_2	χ_3	χ_4
ASP 36(A)	-0.04	0.15	-0.02	-0.24	0.07	0.00	0.00	-0.75	0.34	0.32	-0.79	-0.61	0.00	0.00
LYS 92(A)	-0.39	-0.03	-0.02	-0.3	0.05	0.02	-0.1	0.58	0.36	0.72	-1.22	-0.08	0.17	0.63
ASN 188(A)	0.07	-0.06	-0.06	0.15	0.05	0.00	0.00	0.52	-0.15	0.08	0.18	0.42	0.00	0.00
GLU 45(B)	0.11	-0.15	-0.11	0.16	0.10	0.12	0.00	1.22	-0.01	-0.06	0.28	0.51	0.5	0.00
GLU 65(B)	0.22	0.38	-0.12	-0.28	-0.19	0.43	0.00	-1.76	-0.02	-0.72	-0.16	-1.97	1.12	0.00
ASP 74(B)	-0.54	-0.14	0.25	-0.41	-0.23	0.00	0.00	-2.04	-0.09	0.09	-1.53	-0.50	0.00	0.00
ASP 90(B)	-0.98	-0.12	-0.01	-0.53	-0.32	0.00	0.00	-2.31	-0.11	-0.13	-0.99	-1.07	0.00	0.00
LYS 92(B)	-0.17	-0.11	-0.03	-0.27	0.23	-0.03	0.04	2.25	-0.49	-0.13	-1.93	1.66	0.42	2.72

Table 5. Changes in conformational free energy and entropy (kJ/mol) of the residues in cleavage region in Ca^{2+}(S)-EcoRV-DNA with respect to Ca^{2+}-EcoRV.

Residues	ΔG_i^{conf} (T)	φ	ψ	χ_1	χ_2	χ_3	χ_4	$T\Delta S_i^{conf}$ (T)	φ	Ψ	χ_1	χ_2	χ_3	χ_4
ASP 74(A)	0.20	0.14	0.21	-0.14	-0.00	0.00	0.00	0.59	-0.26	-0.06	0.04	0.88	0.00	0.00
ASP 90(A)	0.96	-0.02	-0.04	0.49	0.52	0.00	0.00	5.09	0.04	-0.28	2.27	3.06	0.00	0.00
LYS 92(A)	0.91	0.21	-0.32	0.45	0.42	-0.03	0.19	5.66	0.29	0.35	2.04	2.65	-0.81	1.14
GLU 45(B)	0.07	0.14	-0.24	0.07	0.09	0.00	0.00	-3.2	-0.22	0.04	0.76	2.38	-6.17	0.00
ASP 74(B)	3.60	-0.11	3.77	0.06	-0.12	0.00	0.00	-1.35	-0.28	-1.83	1.32	-0.56	0.00	0.00
ASP 90(B)	0.86	-0.05	0.03	0.53	0.35	0.00	0.00	4.98	-0.13	-0.03	3.5	1.68	0.00	0.00
LYS 92(B)	0.55	0.06	-0.06	0.4	0.15	0.00	0.00	-6.01	0.16	0.80	1.86	1.08	-6.66	-3.27

The conformational thermodynamics data for the base-pair steps of DNA in Ca^{2+} bound complex with respect to free DNA $\Delta G_{i,i+1}^{fCaS}(b)$ and $T\Delta S_{i,i+1}^{fCaS}(b)$ are shown in Fig. 9(c) and (d). Significant stabilization in free energy occurs only for (4,5) base pair step through ω and Dy (Fig. 9(c)). Both the base pair steps (4,5) and (8,9) are highly ordered through ω, Dx, and Dy (Fig. 9(d)). The changes in free energy and entropy of the residues of EcoRV in cleavage region $\Delta G_i^{conf}(\theta)$ and $T\Delta S_i^{conf}(\theta)$ are shown in Table 5. The most significant change is in destabilization of ASP74(B), the primary contribution being due to ψ. On the other hand, ASP90(A), LYS92(A), and ASP90(B) are disordered through both χ_1 and χ_2. LYS92(B) shows conformational order through χ_3 and χ_4.

Table 6. Total changes (kJ/mol) in conformational thermodynamics in recognition region. We indicate the reference state for each case.

System	Reference state	ΔG_{total}^{DNA}	Rotational	Translational	$T\Delta S_{total}^{DNA}$	Rotational	Translational	ΔG_{total}^{EcoRV}	Back-bone	Side-chain	$T\Delta S_{total}^{EcoRV}$	Back bone	Side-chain	$\Delta G_{Total}^{complex}$	$T\Delta S_{Total}^{complex}$
Mg^{2+}(A) - EcoRV -DNA	Mg^{2+} - EcoRV + free DNA	-24.25	-6.33	-17.92	-34.32	-13.83	-20.49	-12.35	-8.6	-3.75	-24.00	-4.89	-19.11	-36.6	-58.31
Ca^{2+} (S)- EcoRV -DNA	Ca^{2+} - EcoRV+ free DNA	-3.43	-0.90	-2.53	-16.22	-4.1	-12.1	-1.61	-6.04	4.43	14.59	-21.05	35.64	-5.04	-1.63
Ca^{2+}(A)- EcoRV -DNA	Mg^{2+} (A)- EcoRV- DNA	5.61	4.76	0.85	15.56	7.68	7.88	-4.22	-1.15	-2.97	-16.01	-2.5	-13.51	1.39	-0.45
Mg^{2+}(S)- EcoRV -DNA	Ca^{2+} (S) -EcoRV- DNA	-2.98	-2.29	-0.69	-7.31	-2.76	-4.55	3.4	1.22	2.12	-2.55	-0.07	-2.48	0.42	-9.86

ΔG_{total}^{DNA} and $T\Delta S_{total}^{DNA}$ represent the total changes in conformational free energy and entropy respectively in the complex compared to the free states, obtained by adding all the individual contribution of DNA base pair step parameters of all base pair steps in recognition sequence, while ΔG_{total}^{EcoRV} and $T\Delta S_{total}^{EcoRV}$ are those for EcoRV. $\Delta G_{total}^{complex}$ ($\equiv \Delta G_{total}^{DNA} + \Delta G_{total}^{EcoRV}$) and $T\Delta S_{total}^{complex}$ ($\equiv T\Delta S_{total}^{DNA} + T\Delta S_{total}^{EcoRV}$) are the overall changes in conformational free energy and entropy in the complexes with respect to the free components. The conformational thermodynamics data in the recognition region are shown in Table 6 along with the respective reference states and break up of different contributions. The translational parameters give more contributions than rotational parameters in thermodynamic changes of DNA. In case of EcoRV, the backbone dihedrals contribute more to the total free energy, while side chain dihedrals give more contribution to entropy change. Overall, the Mg^{2+}(A)-EcoRV-DNA complex is stabilized

and ordered compared to the free DNA and Mg^{2+}-EcoRV. The major contributions to the changes in free energy and entropy in recognition region come from DNA itself. There is conformational stabilization and order in Ca^{2+}(S)-EcoRV-DNA complex also compared to free DNA and Ca^{2+}-EcoRV. The DNA fragment is stabilized and ordered in the Ca^{2+}(S)-EcoRV-DNA complex. The protein becomes disordered due to enhanced side chain fluctuations unlike that in Mg^{2+}(A)-EcoRV-DNA complex.

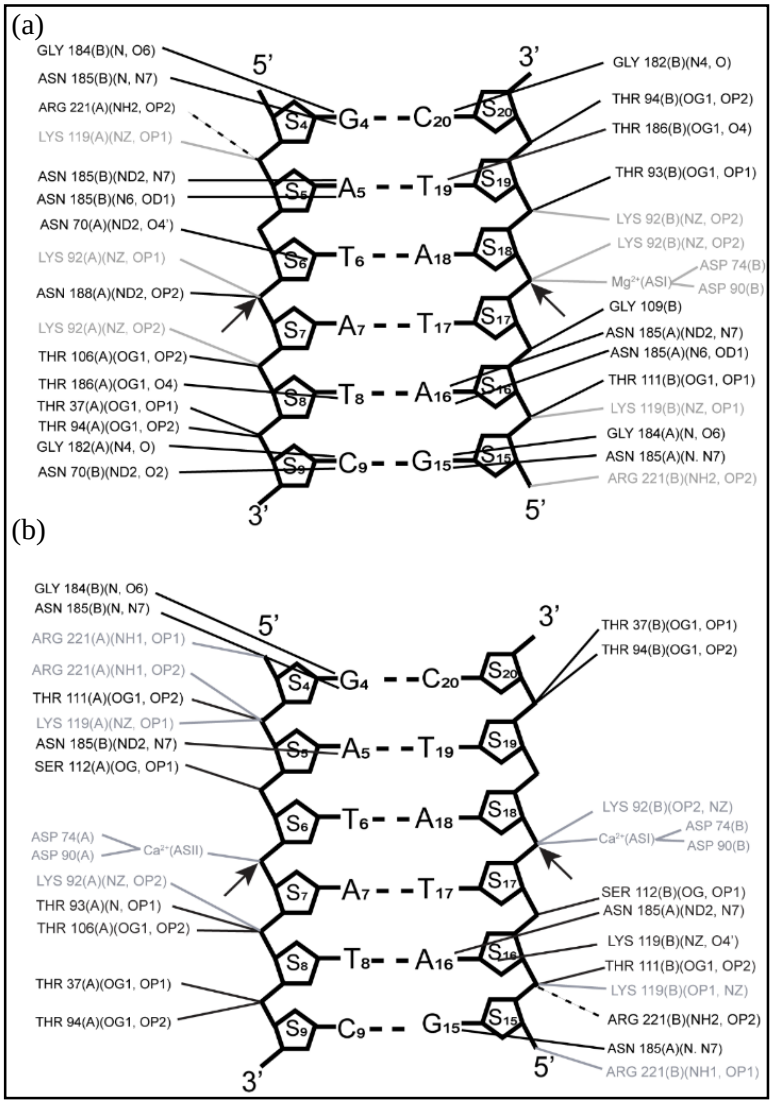

Fig. 10. Interactions between EcoRV and DNA in (a) Mg^{2+}(A)-EcoRV-DNA and (b) Ca^{2+}(S)-EcoRV-DNA. Hydrogen bonding (HB) interactions are illustrated in solid black lines, Electrostatic (EL) interactions in solid grey lines and salt bridge (SB) interactions in dotted black lines. The participating atoms of interaction are shown within brackets.

The conformational stability and order is less in Ca^{2+}(S)-EcoRV-DNA complex than Mg^{2+}(A)-EcoRV-DNA complex, although both complexes are stabilized and ordered, in agreement to experimental data.[12,55] In order to contrast two complexes at microscopic level, we note the differences in their interaction patterns at the protein-DNA interface for various complexes in the cleavage region (see Materials and methods for details). We show hydrogen bond (HB, black solid lines), electrostatic (EL, grey lines) and salt bridge (SB, black dotted lines) interactions in Fig. 10(a) and (b) for Mg^{2+}(A)-EcoRV-DNA and Ca^{2+}(S)-EcoRV-DNA complexes respectively, the participating atoms in the interactions shown within brackets. The total number of HB interactions between protein residues and DNA bases in the recognition region are 13 in Mg^{2+}(A)-EcoRV-DNA complex. Similarly, for sugar and phosphate groups the total number of HB interactions are 1 and 8 respectively. EL interactions occur only between EcoRV residues and the phosphate groups of DNA. There are 7 EL interactions between EcoRV and DNA in the recognition region. The metal ion, Mg^{2+}(ASI) coordinates with EcoRV residues and DNA backbone through EL interaction. There is only one SB interaction. The total number of HB interactions in Ca^{2+}(S)-EcoRV-DNA between EcoRV residues and DNA bases is 5 much less than in Mg^{2+}(A)-EcoRV-DNA complex. The HB interaction between EcoRV residues and sugar and phosphate groups are 1 and 10 respectively. The total number of EL interactions between EcoRV residues and DNA phosphate groups are 7. The number of SB interaction is just 1.

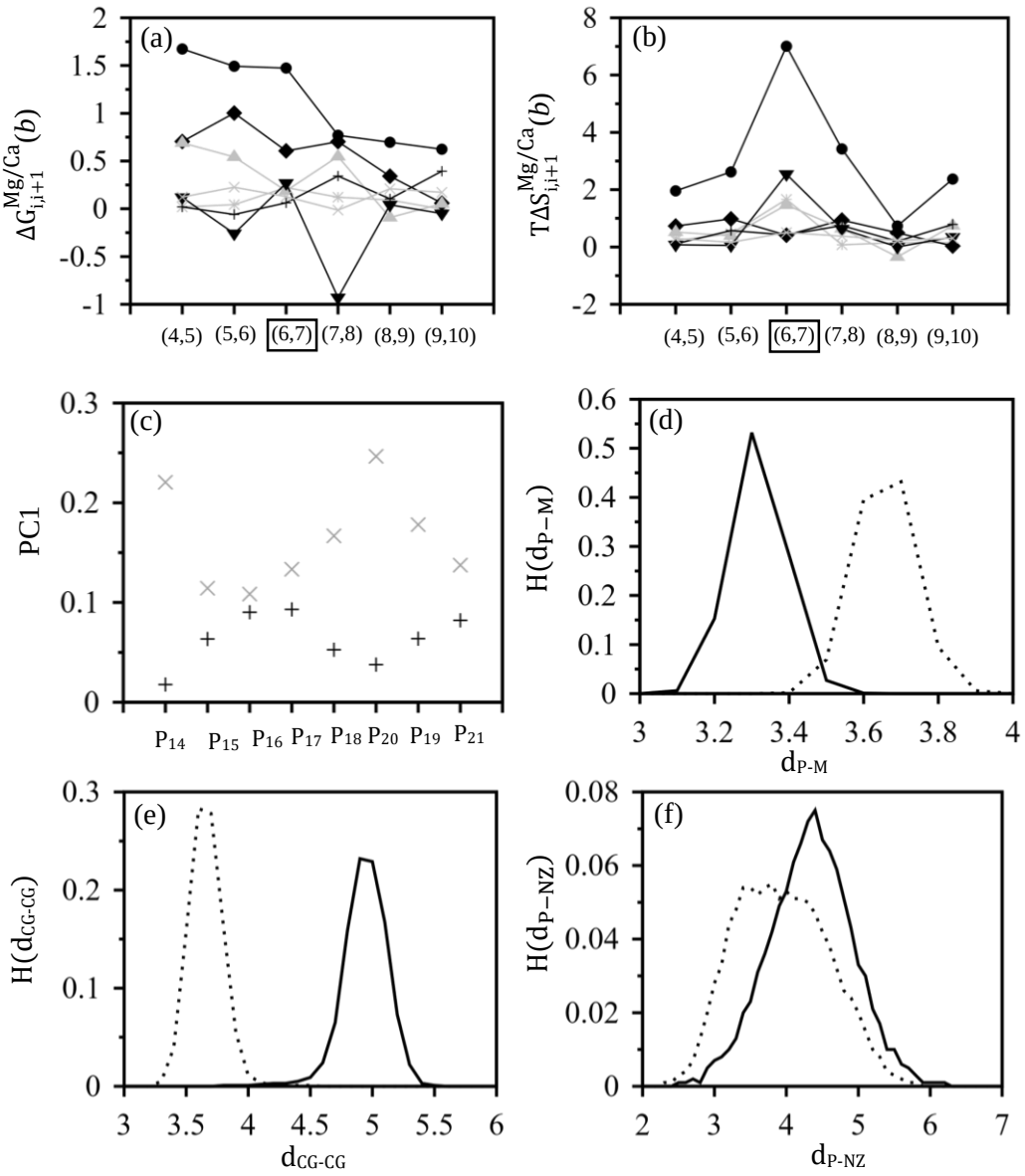

Fig. 11. (a) $\Delta G_{i,i+1}^{Mg/Ca}(b)$ and (b) $T\Delta S_{i,i+1}^{Mg/Ca}(b)$ (in kJ/mol) of DNA base pair step parameters in Ca^{2+}(S)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA. The symbols are: τ (square), ρ (diamond), ω (up triangle), Dx (down triangle), Dy (cross) and Dz (diagonal cross). The total changes are shown in circles. The cleavage base pair has been marked in box. (c) PC1 (first eigenvalue of principal component analysis) of P atoms of the 5'-3' DNA strand in Mg^{2+}(A)-EcoRV-DNA (black cross) and Ca^{2+}(S)-EcoRV-DNA (grey diagonal cross). (d) Distributions $H(d_{P-M})$ of distance d_{P-M} (Å) between P_{18} and Mg^{2+} in Mg^{2+}(A)-EcoRV-DNA (solid line) and Ca^{2+} ion in Ca^{2+}(S)-EcoRV-DNA (dotted line). (e) Distributions $H(d_{CG-CG})$ of the distances d_{CG-CG} (Å) between CG of ASP74(B) and CG of ASP90(B) in Mg^{2+}(A)-EcoRV-DNA (solid line) and Ca^{2+}(S)-EcoRV-DNA (dotted line) complexes. (f) Distributions $H(d_{P-NZ})$ of the distance d_{P-NZ} (Å) between P_{18} and NZ of LYS92(B) in Mg^{2+}(A)-EcoRV-DNA (solid line) and Ca^{2+}(S)-EcoRV-DNA (dotted line).

Table 7. Changes in conformational free energy and entropy (kJ/mol) of the residues in cleavage region in Ca^{2+}(S)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA along with PC1 of selected atoms.

Res	ΔG_i^{conf}	φ	ψ	χ_1	χ_2	χ_3	χ_4	$T\Delta S_i^{conf}$	φ	ψ	χ_1	χ_2	χ_3	χ_4	PC1 for Mg^{2+} - EcoEV-DNA	PC1 for Ca^{2+} - EcoEV-DNA
ASP74(B)	0.06	0.24	-0.14	0.04	-0.09	0.00	0.00	-2.0	-0.01	0.13	0.03	-2.15	0.00	0.00	0.06 (OD2)	0.02 (OD2)
ASP90(B)	2.65	-0.11	-0.16	2.61	0.32	0.00	0.00	2.08	-0.25	-0.42	2.27	0.49	0.00	0.00	0.06 (OD1) 0.06 (OD2)	0.11 (OD1) 0.21 (OD2)
LYS92(B)	-0.07	-0.13	0.03	0.16	-0.14	-0.1	0.12	-1.14	0.00	0.43	0.90	-1.15	-0.82	-0.5	0.02 (CA)	0.01 (CA)

2.3.2 Specificity of metal ion

Let us now concentrate on the differences in the conformational stability and order at microscopic level induced by different metal ions. To this end, the thermodynamic data for Ca^{2+}(S)-EcoRV-DNA complex are computed with respect to Mg^{2+}(A)-EcoRV-DNA complex directly. We consider chain B of EcoRV which holds the metal ions near the cleavage point in both complexes (Mg^{2+}(ASI) and Ca^{2+} (ASI)). The conformational free energy ($\Delta G_{i,i+1}^{Mg/Ca}(b)$) and entropy ($\Delta S_{i,i+1}^{Mg/Ca}(b)$) data for each base pair along with individual base pair step parameter contribution for the DNA base-pair steps are shown in Fig. 11(a) and (b) respectively. It is evident that the base-pair steps are destabilized and disordered in Ca^{2+}(S)-EcoRV-DNA compared to Mg^{2+}(A)-EcoRV-DNA. The (4,5) step gets maximum destabilization. The

maximum conformational disorder is shown by base pair step (6,7) at the cleavage site compared to the Mg^{2+} case. The free energy destabilization is primarily due to the rotational parameters, whereas disorder is derived from all the rotational and the translational parameters. Thus, the cleavage site gets more flexibility in presence of Ca^{2+} than Mg^{2+}. The protein residues in the vicinity of cleavage point also show sensitivity to the metal ion, as can be seen in Table 7, showing $\Delta G_i^{conf}(\theta)$ and $T\Delta S_i^{conf}(\theta)$ data for the residues along with contributions of individual conformational degrees of freedom, comparing the two complexes. ASP74(B) gets ordered, with only marginal change in conformational stability in Ca^{2+}(S)-EcoRV-DNA compared to Mg^{2+}(A)-EcoRV-DNA. Conformational destabilization and disorder in Ca^{2+}(S)-EcoRV-DNA compared to Mg^{2+}(A)-EcoRV-DNA are observed for ASP90(B) primarily through χ_1 fluctuations. We consider LYS92(B) as well. Although this basic residue in the interface of the cleavage region does not participate directly in metal ion coordination, it shows significant changes in conformational thermodynamics in the complex, compared to the free components. LYS92(B) shows order in presence of Ca^{2+} compared to Mg^{2+} case through side chain fluctuations along with marginal stability.

We also have done principal component analysis (PCA) for positional fluctuations in both DNA and EcoRV. We show in Fig. 11(c), the first principal component (PC1) of P of 3'-5' strand. In Mg^{2+}(A)-EcoRV-DNA complex, the fluctuations are having minima at P_{14} and P_{19}. Fluctuation at P_{18} belongs to the basin of minimum at $P_{19.}$ PC1 of P in Ca^{2+}(S)-EcoRV-DNA show contrasting behaviour, having larger fluctuations than in Mg^{2+} bound complex. In particular, the fluctuations show maxima at P_{14} and $P_{19.}$ For DNA base pair step in cleavage region (6,7), PC1 of the center of the mass of base pair step in case of Ca^{2+} complex (0.5) is larger than that for Mg^{2+} complex (0.35). So, larger fluctuation is observed for Ca^{2+} bound complex as compared with Mg^{2+} bound complex, which is in agreement to our thermodynamic data (Fig. 11(b)). We show PC1 of relevant atoms (indicated within brackets) of the protein

residues in the vicinity of cleavage point in chain B in Table 7. Our thermodynamic data, also shown in Table 7, suggest that ASP74(B) is ordered through side chain dihedral in Ca^{2+} bound complex compared with Mg^{2+} bound complex. First principal component (PC1) data show smaller fluctuations of this residue in Ca^{2+}(S)-EcoRV-DNA than in Mg^{2+}(A)-EcoRV-DNA. ASP90(B) shows significant change in conformational thermodynamics. It is destabilized and disordered in Ca^{2+} bound complex when compare with Mg^{2+} bound complex. The disorder of this particular residue comes from the side chain dihedrals. Our PCA data, in same table, also shows greater fluctuations in side chain atoms, OD1 and OD2 in this case, which participate in side chain dihedral fluctuations. We observe that CA atom of LYS92(B), conformationally ordered in Ca^{2+}(S)-EcoRV-DNA complex, shows more fluctuations in Mg^{2+}(A)-EcoRV-DNA complex than in Ca^{2+}(S)-EcoRV-DNA complex. So, the PCA data are consistent with conformational thermodynamics data for protein residues and DNA base pair steps.

The differences in conformational stability and order at the cleavage region, having no HB interaction, can be understood in terms of electrostatic interactions. We show in Fig. 11(d), histograms $H(d_{P\text{-}M})$ of the distances ($d_{P\text{-}M}$) between oxygen atom (OP1) of scissile phosphate P_{18} and Mg^{2+}(ASI), and Ca^{2+}(ASI). We observe that Mg^{2+}(ASI) is closer to the phosphate group in Mg^{2+} bound complex than Ca^{2+}(ASI). So, the electrostatic attraction between Mg^{2+} and phosphate group is stronger than in Ca^{2+} bound case. The DNA base pair steps have more rotational flexibility in presence of Ca^{2+} than that in presence of Mg^{2+} due to weaker electrostatic interactions.

Next, we consider the protein residues. While ASP90(B) holds Mg^{2+}(ASI) in the vicinity of P_{18}, ASP74(B) is coordinated to both Mg^{2+}(ASI) and Mg^{2+}(ASII). But for Ca^{2+} bound complex, both ASP74(B) and ASP90(B) hold Ca^{2+}(ASI) in the vicinity of P_{18}. We show the histograms, $H(d_{CG\text{-}CG})$ of distance ($d_{CG\text{-}CG}$) between CG atoms in side chains of ASP74(B) and ASP90(B) in Fig. 11(e), for both Mg^{2+} and Ca^{2+} bound complexes. Solid lines represent the

case of Mg^{2+} bound complex, while dotted lines for Ca^{2+} bound complex. The side chains of two aspartic acids are closer in Ca^{2+} bound complex than in Mg^{2+} bound complex. This is because both the residues coordinate with one metal ion in Ca^{2+} bound complex. Due to similar charged states, these residues experience enhanced repulsion, leading to enhanced side- chain fluctuations in ASP90(B). LYS92(B), on the other hand, shows marginal conformational stability and order in Ca^{2+} complex than Mg^{2+} complex, indicating role in structural stabilization in Ca^{2+} complex. We calculate histograms $H(d_{P\text{-}NZ})$ of the distance ($d_{P\text{-}NZ}$) between phosphorus of P_{18} and positively charged NZ atom in side chain of LYS92(B) (Fig. 11(f)). We find that NZ atom is closer to P_{18} but with larger fluctuations in Ca^{2+} complex than in Mg^{2+} complex. This suggests competitive electrostatic interactions of NZ atom with metal ion and P_{18}. Weaker electrostatic interaction between metal ion and P_{18}, helps NZ atom to approach P_{18}, closer than in Mg^{2+}(A)-EcoRV-DNA complex. This leads to conformational stability of the residue.

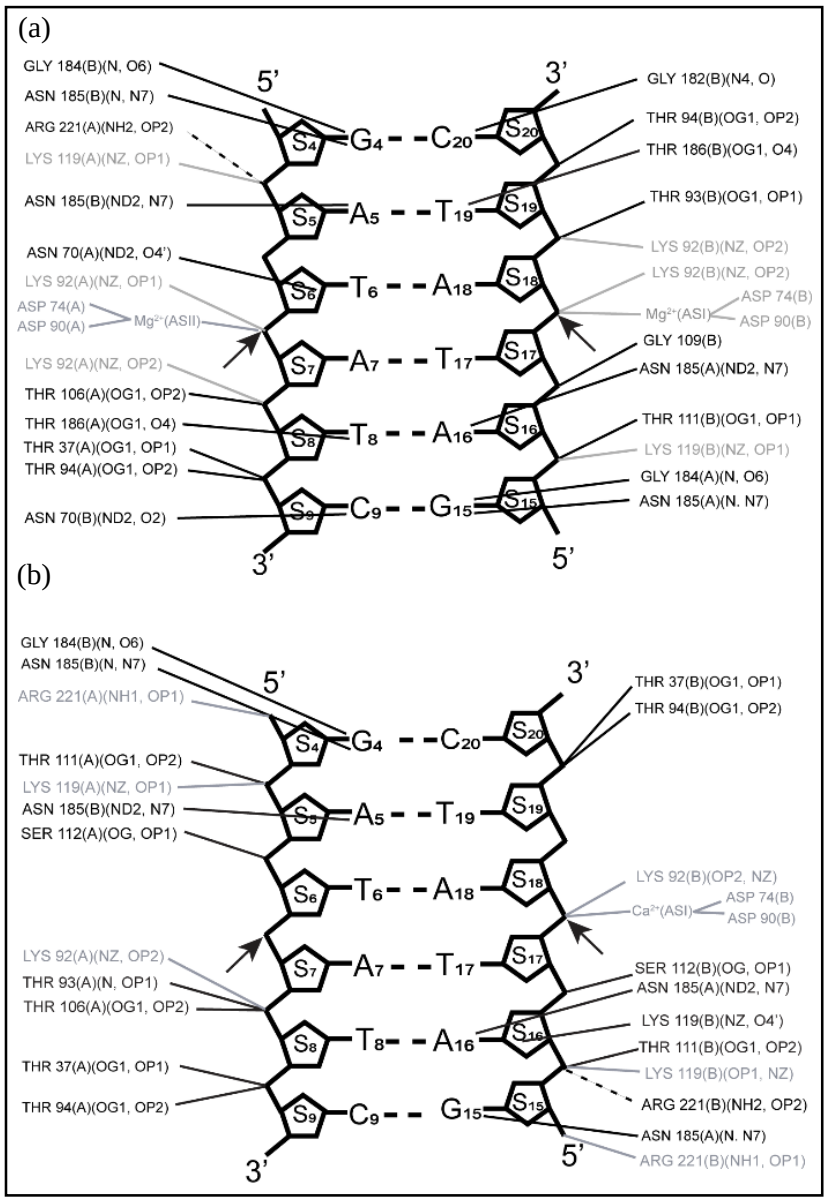

Fig. 12. Interactions between EcoRV and DNA in (a) Mg^{2+}(S)-EcoRV-DNA and (b) Ca^{2+}(A)-EcoRV-DNA. Hydrogen bonding (HB) interactions are illustrated in solid black lines, Electrostatic (EL) interactions in solid grey lines and salt bridge (SB) interactions in dotted black lines. The participating atoms of interaction are shown within brackets.

Now we turn our attention to the specificity to the identity of the metal ions considering a couple of model systems, namely, Mg^{2+} at symmetric locations in place of Ca^{2+}, denoted by Mg^{2+}(S)-EcoRV-DNA) and Ca^{2+} in place of Mg^{2+}at asymmetric locations, denoted by Ca^{2+}(A)-EcoRV-DNA. Since we observe that 50 ns time windows for Mg^{2+}(A)-EcoRV-DNA and Ca^{2+}(S)-EcoRV-DNA complexes yield well convergent thermodynamic quantities, we restrict

to shorter runs for the model complexes. The microscopic interactions at the protein-DNA interfaces are shown for Mg^{2+}(S)-EcoRV-DNA and Ca^{2+}(A)-EcoRV-DNA complexes in Fig. 12(a) and (b). The total number of HB interactions between EcoRV and the DNA bases are 10 in Mg^{2+}(S)-EcoRV-DNA, while this number is 5 in Ca^{2+}(A)-EcoRV-DNA. There are 1 and 7 HB interactions between EcoRV and sugar and phosphate groups in Mg^{2+}(S)-EcoRV-DNA. Similarly, the total number of HB interactions between EcoRV and phosphate groups in Ca^{2+}(A)-EcoRV-DNA are 10. There is 1 HB interaction between EcoRV and the sugar group in Ca^{2+}(A)-EcoRV-DNA. The total number of EL and SB interactions between EcoRV residues and the phosphate groups are 6 and 1 in both the cases. Both Mg^{2+} ions in Mg^{2+}(S)-EcoRV-DNA and Ca^{2+}(ASI) ion in Ca^{2+}(A)-EcoRV-DNA participate in EL interactions with the phosphate groups of DNA. Thus, comparing the data for different complexes (Figs. 10(a) and (b), and Figs. 12(a) and (b)), it is clear that the microscopic interaction pattern is sensitive to both the type and locations of the metal ions.

The total free energy and entropy changes in recognition region for Ca^{2+}(A)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA and Mg^{2+}(S)-EcoRV-DNA with respect to Ca^{2+}(S)-EcoRV-DNA are shown in Table 6. In Ca^{2+}(A)-EcoRV-DNA complex, DNA is destabilized through rotational parameters and disordered through both rotational and translational parameters. The destabilization and disorder are counterbalanced by stabilization and order in the protein residues mainly via the side chain fluctuations. In case of Mg^{2+}(S)-EcoRV-DNA, the DNA base pairs are stabilized but the protein residues undergo destabilization, while both show marginal order. Overall, the model complexes show marginal changes in conformational thermodynamics.

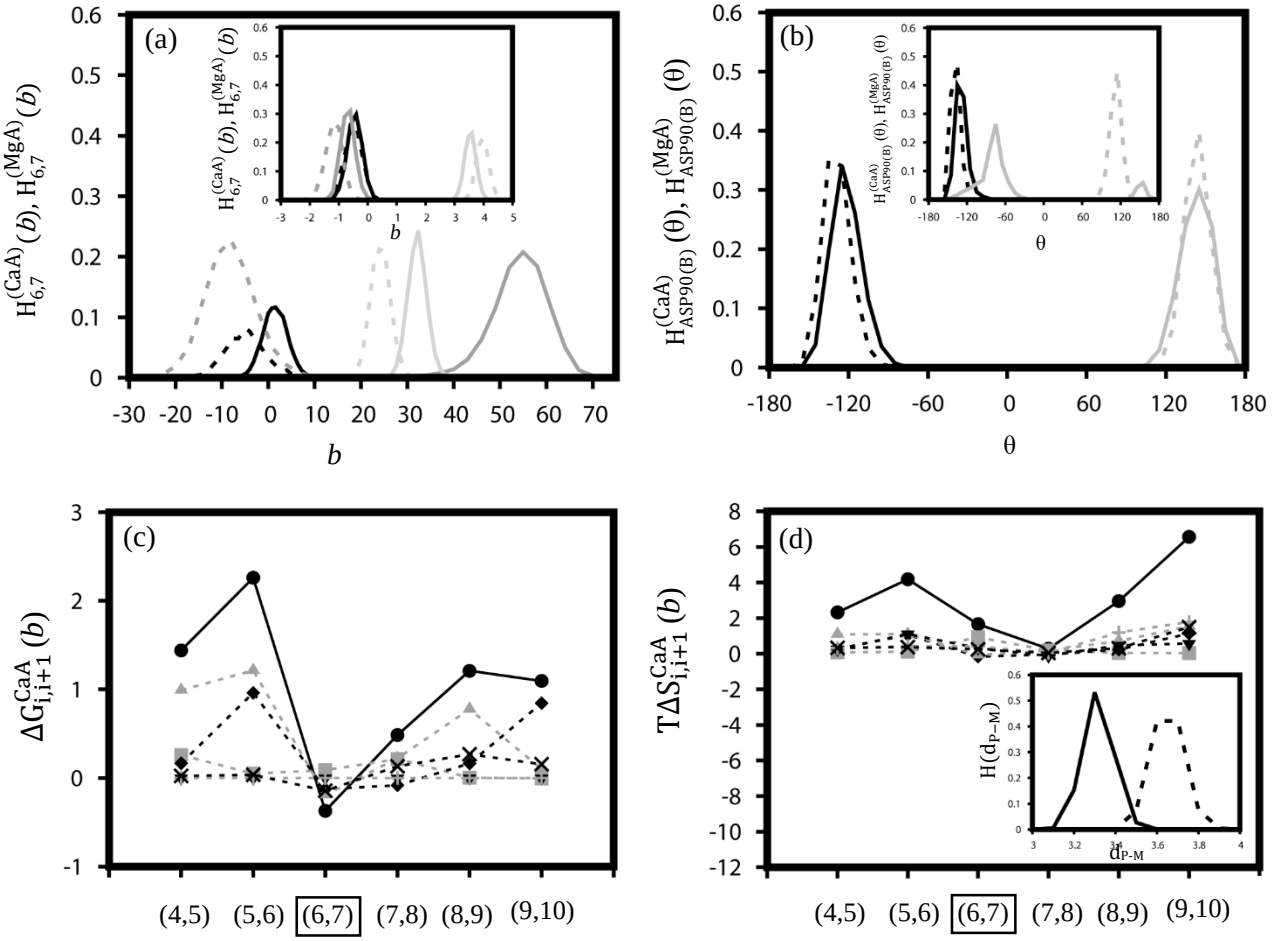

Fig. 13. (a) $H_{i,i+1}^{(CaA)}(b)$ and $H_{i,i+1}^{(MgA)}(b)$ of the rotational parameters of the DNA base pair step (6,7) in Mg^{2+}(A)-EcoRV-DNA (solid line) and Ca^{2+}(A)-EcoRV-DNA (dotted line); (b) $H_{i}^{(CaA)}(\theta)$ and $H_{i}^{(MgA)}(\theta)$ of backbone dihedral angles of the residue ASP90(B) in Mg^{2+}(A)-EcoRV-DNA (solid line) and Ca^{2+}(A)-EcoRV-DNA (dotted line). The inset in (a) illustrates the distributions of translational parameters of the same base pair step as in (a). Similarly, the inset in (b) illustrates the distributions of side chain dihedrals of the same residue as in (b). The color conventions are: τ (black), ρ (grey), ω (light grey), Dx (black), Dy (grey) and Dz (light grey) for DNA base pair step parameters, and φ (black), ψ (grey), χ_1 (black) and χ_2 (grey) for protein dihedral angles. (c) $\Delta G_{i,i+1}^{CaA}(b)$ and (d) $T\Delta S_{i,i+1}^{CaA}(b)$ of base pair step parameters of DNA base pair steps (in kJ/mol) in Ca^{2+}(A)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA. Here, τ (square), ρ (diamond), ω (up triangle), Dx (down triangle), Dy (cross) and Dz (diagonal cross). The cleavage base pair step has been marked in box. Inset (d) shows the distributions H ($d_{P\text{-}M}$) of distance $d_{P\text{-}M}$ (Å) between OP1 of P_{18} and Mg^{2+} in Mg^{2+}(A)-EcoRV-DNA (solid line) and Ca^{2+} in Ca^{2+}(A)-EcoRV-DNA (dotted line).

Let us consider the microscopic details of the Ca^{2+}(A)-EcoRV-DNA system first. We do not observe any rearrangement of coordination of the metal ions in the complex compared to Mg^{2+}(A)-EcoRV-DNA complex. Representative histograms of base pair step parameters (denoted by $H_{i,i+1}^{(CaA)}(b)$) and dihedral angles (denoted by $H_{i}^{(CaA)}(\theta)$) in the cleavage region and

similar data for Mg^{2+}(A)-EcoRV-DNA complex (denoted by $H_{i,i+1}^{(MgA)}(b)$ and $H_{i}^{(MgA)}(\theta)$) are shown in Fig. 13 (a) and (b) respectively. Let us highlight the contrasts. The histograms of τ and ρ of the base pair step (6,7) are single peaked in both complexes, but broader in Mg^{2+}(S) - EcoRV-DNA compared to the other complex, while ω is sharp single peaked in both cases (Fig. 13(a)). The distributions of Dx, Dy and Dz are all single peaked in both cases, shown in inset of Fig. 13(a). Next, we consider the fluctuations in dihedrals of the residue ASP90(B) in cleavage region. The backbone dihedrals show sharp single peak in both cases (Fig. 13(b)). χ_1 in ASP90(B) is single peaked in both cases (see inset of Fig. 13(b)). Similarly, χ_2 is multimodal distribution in Mg^{2+}(A)-EcoRV-DNA, while single peaked in Ca^{2+}(A)-EcoRV-DNA (see inset of Fig. 13(b)).

Table 8. The changes in conformational thermodynamics of metal ion (kJ/mol) coordinating EcoRV residues in cleavage region of Ca^{2+}(A)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA.

Residues	ΔG_i^{conf} (T)	φ	ψ	χ_1	χ_2	χ_3	χ_4	$T\Delta S_i^{conf}$ (T)	φ	Ψ	χ_1	χ_2	χ_3	χ_4
ASP36(A)	-0.79	-0.40	-0.25	-0.06	-0.08	0.00	0.00	-0.63	-0.61	-0.58	0.41	0.16	0.00	0.00
GLU45(B)	-1.7	-0.10	0.01	-0.90	-0.44	-0.29	0.00	-7.94	-0.13	-0.24	-2.65	-3.18	-1.75	0.00
ASP74(B)	-0.69	0.16	-0.29	-.012	-0.43	0.00	0.00	-3.77	-0.02	-0.09	-0.56	-3.09	0.00	0.00
ASP90(B)	-1.04	-0.03	-0.25	-0.39	-0.37	0.00	0.00	-3.67	-0.39	-0.44	-0.65	-2.2	0.00	0.00

The thermodynamic changes of the EcoRV residues ($\Delta G_i^{conf}(\theta)$ and $T\Delta S_i^{conf}(\theta)$)) in cleavage region in Ca^{2+}(A)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA are shown in Table 8. The protein residues are marginally stabilized and ordered through the side chain fluctuations. This is unlike the case observed in Ca^{2+}(S)-EcoRV-DNA complex compared to Mg^{2+}(A)-EcoRV-DNA in Table 7. However, the DNA base pair step parameters show destabilization and disorder in Ca^{2+}(A)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA, similar to those observed in Ca^{2+}(S)-EcoRV-DNA compared to Mg^{2+}(A)-EcoRV-DNA complex (Fig. 11 (a) and (b)). The details of changes in conformational free energy and entropy

of the DNA base pair step parameters (denoted by $\Delta G_{i,i+1}^{CaA}(b)$ and $T\Delta S_{i,i+1}^{CaA}(b)$ respectively) of the base pair steps are shown in Fig. 13(c) and (d). The base pair step (5,6) in the vicinity of the cleavage base pair, gets large destabilization through ρ and ω and disorder primarily through ρ, ω and Dx. Inset, Fig. 13(d) shows the histograms of the distances between the oxygen of P_{18} and the metal ions Mg^{2+} in Mg^{2+}(A)-EcoRV-DNA (black solid line) and Ca^{2+} in Ca^{2+}(A)-EcoRV-DNA (black dotted line). The data show that the replacement of Mg^{2+} by Ca^{2+} increases the distance between scissile phosphate group and the metal ion. The increased distance is due to larger ionic size and differences in the LJ parameters for Ca^{2+} ion compared to Mg^{2+} ion which results in less electrostatic interactions in the corresponding complex.

Let us consider microscopic picture in the other model system, Mg^{2+}(S)-EcoRV-DNA. In this complex we observe a slight change in the coordination where GLU 45(A) in addition to the other residues is located within the coordination distance in the second site. Representative distributions of base pair step parameters of DNA base pair steps ($H_{i,i+1}^{(MgS)}(b)$) and EcoRV residues ($H_{i}^{(MgS)}(\theta)$) in cleavage region of Mg^{2+}(S)-EcoRV-DNA along with those in Ca^{2+}(S)-EcoRV-DNA ($H_{i,i+1}^{(CaS)}(b)$ and $H_{i}^{(CaS)}(\theta)$ respectively) are shown in Figs. 14(a) and (b) respectively. τ of (6,7) shows multimodal distribution in Mg^{2+}(S)-EcoRV-DNA, while broad single peak in Ca^{2+}(S)-EcoRV-DNA (Fig. 14(a)). ρ shows single peak with long tail, while ω is broad single peaked in both cases (Fig. 14(a)). Dx shows multimodal distribution in Ca^{2+}(S)-EcoRV-DNA, while single peak in Mg^{2+}(S)-EcoRV-DNA (see inset of Fig. 14(a)). The other translational parameters are single peaked in both cases. The backbone dihedral angles of ASP90(B) (Fig. 14(b)) are all sharp single peaked for both cases. Both χ_1 and χ_2 of ASP90(B) are multimodal distributions in Ca^{2+}(S)-EcoRV-DNA (see inset of Fig. 14(b)). Similarly, χ_1 shows bimodal and χ_2, single peaked distributions in Mg^{2+}(S)-EcoRV-DNA complex (see inset of Fig. 14(b)).

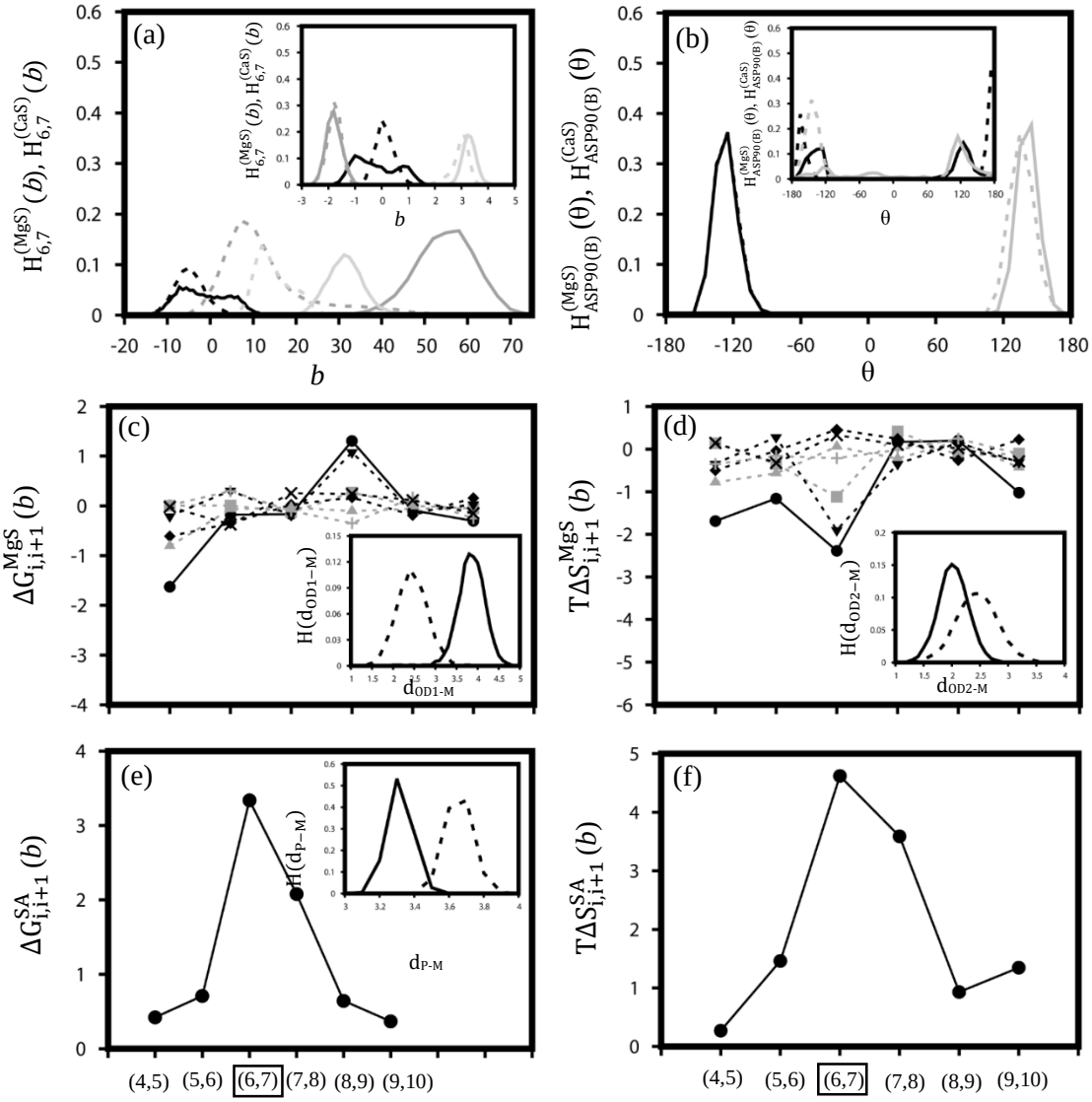

Fig. 14. (a) $H_{i,i+1}^{(MgS)}(b)$ and $H_{i,i+1}^{(CaS)}(b)$ of τ (black), ρ (grey) and ω (light grey) of the DNA base pair step (T_6:A_{18}, A_7:T_{17}) Mg^{2+}(S)-EcoRV-DNA (solid line) and Ca^{2+}(S)-EcoRV-DNA (dotted line); (b) $H_i^{(MgS)}(\theta)$ and $H_i^{(CaS)}(\theta)$ of φ (black) and ψ (light grey) of the residue ASP90(B) in Mg^{2+}(S)-EcoRV-DNA (solid line) and Ca^{2+}(S)-EcoRV-DNA (dotted line). The inset in (a) illustrates the distributions of Dx (black), Dy (grey) and Dz (light grey) of the same base pair step as in (a). Similarly, the inset in (b) illustrates the distributions of χ_1 (black) and χ_2 (grey) of the same residue as in (b). (c) $\Delta G_{i,i+1}^{MgS}(b)$ and (d) $T\Delta S_{i,i+1}^{MgS}(b)$ of base pair step parameters of DNA base pair steps (in kJ/mol) in Mg^{2+}(S)-EcoRV-DNA with respect to Ca^{2+}(S)-EcoRV-DNA. The symbols are: τ (square), ρ (diamond), ω (up triangle), Dx (down triangle), Dy (cross) and Dz (diagonal cross). The cleavage base pair has been marked in box. The inset in (c) shows the distributions $H(d_{OD1\text{-}M})$ of distance $d_{OD1\text{-}M}$ (Å) between OD1 of ASP90B and Ca^{2+}(ASI) in Ca^{2+}(S)-EcoRV-DNA (dotted line) and Mg^{2+}(ASI) in Mg^{2+}(S)-EcoRV-DNA (solid line). Similarly, the inset in (d) illustrates the distributions $H(d_{OD2\text{-}M})$ of distance $d_{OD2\text{-}M}$ (Å) between OD2 of ASP90(B) and Ca^{2+}(ASI) in Ca^{2+}(S)-EcoRV-DNA (dotted line) and Mg^{2+}(ASI) in Mg^{2+}(S)-EcoRV-DNA (solid line). The total change in (e) $\Delta G_{i,i+1}^{SA}(b)$ and (f) $T\Delta S_{i,i+1}^{SA}(b)$ of the base pair step parameters of the base pair steps (in kJ/mol) in Mg^{2+}(S)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA. The inset in (e) shows the histogram H ($d_{P\text{-}M}$) of distance $d_{P\text{-}M}$ (Å) between P_{18} and Mg^{2+}(ASI) ion in Mg^{2+}(A)-EcoRV-DNA (solid line) and Mg^{2+}(ASI) in Mg^{2+}(S)-EcoRV-DNA (dotted line).

Among the protein residues ($\Delta G_i^{conf}(\theta)$ and $T\Delta S_i^{conf}(\theta)$ in Table 9) in the cleavage region, ASP90(A) and ASP90(B), holding the metal ions, undergo destabilization in Mg^{2+}(S)-EcoRV-DNA with respect to Ca^{2+}(S)-EcoRV-DNA complex through side chain fluctuations. This is similar to destabilization observed in this residue in Ca^{2+}(S)-EcoRV-DNA compared to Mg^{2+}(A)-EcoRV-DNA in Table 7, and unlike the observations in Table 8. The destabilization in ASP (90) could result from proximity of ASP (74) residue in both the chains due to smaller ionic size of Mg^{2+}. The changes in conformational thermodynamics of the base pair step parameters of DNA in Mg^{2+}(S)-EcoRV-DNA with respect to Ca^{2+}(S)-EcoRV-DNA ($\Delta G_{i,i+1}^{MgS}(b)$ and $T\Delta S_{i,i+1}^{MgS}(b)$) are shown in Fig. 14 (c) and (d). The base pair step parameters show overall marginal stabilization and order which is unlike the previous case (Fig. 13(c) and (d)). The histogram H($d_{OD1\text{-}M}$) of distance $d_{OD1\text{-}M}$ (Å) between OD1 and the histogram H($d_{OD2\text{-}M}$) of distance $d_{OD2\text{-}M}$ (Å) between OD2 of ASP90B in insets of Fig.14(c) and Fig. 14(d) show that the metal ion is better coordinated to the acidic residue in Ca^{2+}(S)-EcoRV-DNA than in Mg^{2+}(S)-EcoRV-DNA through OD1, while that through OD2 does not change much.

Table 9. The conformational thermodynamic changes of the metal ion coordinating residues (kJ/mol) in cleavage region of Mg^{2+}(S)-EcoRV-DNA with respect to Ca^{2+}(S)-EcoRV-DNA.

Residues	ΔG_i^{conf}	φ	ψ	χ_1	χ_2	χ_3	χ_4	$T\Delta S_i^{conf}$	φ	Ψ	χ_1	χ_2	χ_3	χ_4
GLU45(A)	0.36	-0.06	0.23	-0.28	0.53	-0.06	0.00	-3.24	-0.03	0.48	-2.68	-0.32	-0.69	0.00
ASP74(A)	-0.08	0.09	0.07	-0.37	0.13	0.00	0.00	0.86	0.11	0.37	-1.25	1.62	0.00	0.00
ASP90(A)	2.30	-0.25	1.47	0.57	0.51	0.00	0.00	2.33	-0.25	-0.55	2.75	0.38	0.00	0.00
ASP74(B)	-0.88	-0.29	-0.08	-0.02	-0.48	0.00	0.00	-2.10	-0.35	-0.05	0.75	-2.45	0.00	0.00
ASP90(B)	1.7	-0.13	0.21	1.56	0.07	0.00	0.00	-0.40	-0.16	0.36	-1.36	0.61	0.00	0.00

This also makes us compare between complexes with the symmetric versus asymmetric locations of Mg^{2+}. We compare conformational thermodynamics of the symmetric complex with respect to location of the Mg^{2+} ions (Mg^{2+}(S)-EcoRV-DNA) to that with asymmetrically located metal ions (Mg^{2+}(A)-EcoRV-DNA). The data for changes in conformational free

energy and entropy, denoted by $\Delta G^{SA}_{i,i+1}(b)$ and $T\Delta S^{SA}_{i,i+1}(b)$, shown in Fig.14(e) and (f) respectively, for the base pair steps show the conformational destabilization and disorder for symmetrical arrangement of the metal ions with respect to their asymmetric location. The conformational stability and order of the protein residues in Table 10 do not show any qualitative difference compared to those in Table 9. For instance, ASP90(B) remains unstable in the symmetric complex with respect to the asymmetric one. The enhanced fluctuations in DNA base pairs are reflected in the inset of Fig. 14(e) with larger distance and fluctuations of distance $d_{P\text{-}M}$. Thus, the symmetric complex with Mg^{2+} is less efficient than the asymmetric one in holding the metal ion in the neighbourhood of the scissile phosphate group.

Table 10. Conformational thermodynamic changes of the residues in cleavage region (kJ/mol) in Mg^{2+}(S)-EcoRV-DNA with respect to Mg^{2+}(A)-EcoRV-DNA.

Residues	ΔG_i^{conf}	φ	ψ	χ_1	χ_2	χ_3	χ_4	$T\Delta S_i^{conf}$	φ	Ψ	χ_1	χ_2	χ_3	χ_4
ASP36 (A)	-0.14	-0.11	-0.06	0.09	-0.07	0.00	0.00	0.78	1.04	-1.39	0.81	0.31	0.00	0.00
ASP74 (A)	-0.45	0.14	-0.31	-0.19	-0.09	0.00	0.00	-0.67	0.07	-0.38	-0.78	0.41	0.00	0.00
ASP90 (A)	1.76	-0.18	1.42	0.46	0.07	0.00	0.00	-0.15	-0.33	-0.95	2.36	-1.23	0.00	0.00
GLU45 (B)	-1.16	-0.06	0.01	-0.54	-0.38	-0.19	0.00	-4.74	-0.05	-0.12	-1.16	-2.21	-1.20	0.00
ASP74 (B)	-0.56	0.05	-0.28	0.01	-0.35	0.00	0.00	-1.85	0.09	-0.05	0.52	-2.41	0.00	0.00
ASP90 (B)	0.99	-0.17	-0.02	1.39	-0.22	0.00	0.00	-3.27	-0.26	0.01	-1.59	-1.48	0.00	0.00

Our classical force field-based simulations cannot yield detailed picture of transition states involving charge transfer in the cleavage region. However, the conformational thermodynamics data give useful microscopic insight, if we assume that the metal ion must be stabilized in the vicinity of the scissile phosphate group for cleavage activity. Larger conformational stability and order at the (4,5) step suggests that this particular step helps to stabilize and order the complex. The stability and order at the other base pair steps of the DNA in the cleavage region points to co-operative effect of the entire region. In the stable complex,

the Mg^{2+}(ASI) can access the cleavage phosphate which is further helped by the conformational order in ASP74(B) and ASP90(B) holding the metal ion. Our results are consistent to the experimental observations that catalytic activity is lost upon mutation of ASP74(B) and ASP90(B) [25–27]. Although Mg^{2+}(ASII) does not participate directly in interacting with P_{18}, it helps to hold Mg^{2+}(ASI) in the vicinity of P_{18} by imparting stability to the aspartic acid residues. The role of the second Mg^{2+} has been discussed earlier. Site-directed mutagenesis experiments show that a water molecule, which is a part of the inner hydration sphere of second Mg^{2+}, protonates the leaving group (3'-OH) through a favourable six-membered cycle.[29] Theoretical study by Zahran et. al. shows that the second magnesium ion serves to stabilize the distorted DNA.[14] Our observations on the stability of the aspartic acids is qualitatively consistent to this theoretical report. The larger fluctuations at ASP90 (B) in the Ca^{2+}(S)-EcoRV-DNA complex cannot hold the metal ion in the vicinity of $P_{18.}$ As a result, P_{18} shows larger fluctuations which stands in the way of the cleavage activity in the complex. It may be interesting to point out the role of basic LYS92 (B) in the EcoRV interface with the cleavage region. This residue is not directly involved in coordination of the metal ion. LYS92(B) shows conformational destabilization and disorder in Mg^{2+}(A)-EcoRV-DNA complex compared to the free components, ruling out its role in stabilization in the complex. Earlier studies show [38–41] that the residues which get conformationally destabilized and disordered participate in ligand binding. Thus, in this complex, LYS92(B) is likely to play functional role. The important point here is that the DNA fragment and the protein residues in the cleavage region need to be stabilized and the metal ion is to be placed in the vicinity of the scissile phosphate.

Although the conformational destabilization is the least at the cleavage base pair steps in Ca^{2+}(S)-EcoRV-DNA, larger fluctuations in the neighbouring regions do not hold the metal ion in the vicinity of the scissile phosphate. This further supports the co-operative participation of the DNA base pairs in the cleavage process. Similarly, the fluctuations of DNA base pairs

make it unfavourable in symmetric complex compared to the one with asymmetrically positioned Mg^{2+}, ruling out the possibility of enzymatic activities in the symmetric complex. Our results suggest that two Mg^{2+} ions located on one protein subunit is essential for cleaving one phosphate group. In catalytically active complex cleaving occurs at two scissile phosphate groups on two DNA strands in the cleavage region. This requires four Mg^{2+} ions, two each in two protein subunits to achieve adequate conformational stability and order for cleavage.

2.4 Conclusions

To summarize, microscopic fluctuation-based picture shows that the cleavage region gets conformational stability and order and the metal ion is placed stably in the neighbourhood of the scissile phosphate group in Mg^{2+}(A)-EcoRV-DNA in preference to Ca^{2+}(S)-EcoRV-DNA. On the other hand, replacements of the metal ions in these complexes renders the cleavage region unstable and disordered. This suggests that both the ion type and their locations are decisive for cofactor sensitivity for cleavage activity by EcoRV. More importantly, our proposed scenario based on the fluctuations in microscopic conformational variables can be verified in NMR experiments. The motif PD…(D/E)XK (P – proline, D – aspartic acid, E – glutamic acid, K – Lysine, and X – any amino acid) is identified as common catalytic sequence motif for all type II REs,[8,56,57] where two carboxylates are responsible for Mg^{2+} binding. The activity of this catalytic motif is Mg^{2+} dependent.[58] Not only type II REs, but this catalytic motif is seen in some other REs also, for example, in λ-exonuclease ($PD^{119}…E^{129}LK$),[59] and E. coli MutH ($QD^{70}…E^{77}LK$).[60] The scenario based on the microscopic fluctuations is expected to hold in all such cases. Microscopic fluctuations of the conformational variables, being robust in nature, may provide useful to understand other enzyme complexes and provide in-silico guideline in designing genetic tools for targeted DNA manipulations.

2.5 Appendix A

Details of the methods

A.1 Molecular dynamics simulations

Molecular dynamics (MD) simulation is the study of time dependent behavior of system of particles. MD simulations begin with particles' initial positions and velocities and then an algorithm is used to update each particles positions and velocities. The result is a trajectory of how the particles' positions and velocities change with time. The particles are considered to interact with each other via a pair potential V (r_{ij}), where $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$, $i \neq j$ and $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ are the positions of the i^{th} and j^{th} particles. The trajectory of the particles is obtained by solving Newton's law, $\boldsymbol{F}_i = m_i \boldsymbol{a}_i$, where $\boldsymbol{F}_i$ is the force acting on i^{th} particle, m_i is the mass and $\boldsymbol{a}_i$ is the acceleration of the i^{th} particle. $\boldsymbol{F}_i$ can be obtained from the gradient of V (r_{ij}). The most widely used algorithm is Verlet algorithm which is based on central difference algorithm. In Verlet algorithm, the advanced position $\vec{r}(t + \Delta t)$, at time $t + \Delta t$ is calculated by using the positions and the accelerations at time t and the positions from the previous step, $\vec{r}(t - \Delta t)$. Using Taylor series expansion,

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \Delta t\, \vec{v}(t) + \frac{\Delta t^2}{2} \vec{a}(t) + \cdots \quad (1)$$

$$\vec{r}(t - \Delta t) = \vec{r}(t) - \Delta t\, \vec{v}(t) + \frac{\Delta t^2}{2} \vec{a}(t) + \cdots \quad (2)$$

Combined these two equations gives,

$$\vec{r}(t + \Delta t) = 2\vec{r}(t) - \vec{r}(t - \Delta t) + \Delta t^2 \vec{a}(t) \quad (3)$$

The velocity does not appear explicitly in Verlet algorithm. It can be obtained from the following formula, $\vec{v}(t) = \frac{\vec{r}(t+\Delta t) - \vec{r}(t-\Delta t)}{2\Delta t}$. (4)

A.2 Force-field for MD simulations

The packages we have used in our study to simulate biomolecular complexes are NAMD and GROMACS. These standard packages use force field to describe the interactions between atoms within the biomolecules or biomolecules with another biomolecules. The parameters of a force field are often derived through ab initio or semi-empirical quantum mechanical calculations, or by fitting experimental data such as neutron, X-ray and electron diffraction, NMR etc. The appropriate choice of force field (CHARMM, GROMOS and AMBER) is necessary in order to improve the accuracy of the results. All common force fields contain bonded and non-bonded interactions. The bonded interactions include bond stretching, bond rotations and torsional dihedrals via simple harmonic oscillations, whereas non-bonded interactions include van der Waals and Coulomb interactions.

The expression of a potential energy is given by,

$$U = \sum_{bonds} k_b(r - r_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{torsions} k_\varphi\,[1 + \cos(n\varphi - \delta)] + \sum_{impropers} k_w(w - w_0)^2 + \sum_{LJ} 4\varepsilon\,[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^{6}] + \sum_{elec} \frac{q_i q_j}{4\pi\varepsilon_r\varepsilon_0 r_{ij}} \quad (5)$$

Here, the first four terms refer to the contributions to the total energy due to intramolecular interactions, and the last two terms refer to the contribution to total energy due to non-bonded interactions. k_b is force constant due to bond stretching, r_0 is equilibrium bond length, while $(r - r_0)$ is the deviation of the bond length. Similarly, $(\theta - \theta_0)$ is the deviation of the angle from an equilibrium angle θ_0 with force constant k_θ. k_φ is dihedral force constant, n is the multiplicity number and δ is phase shift. $(w - w_0)$ is the deviation of the outer plane angle from w_0 with force constant k_w. For non-bonded interactions, ε is the depth of the potential, σ_{ij} implies the distance at which the intermolecular potential between two particles is zero. The

last term represents the Coulomb interactions, where q_i and q_j are the charges of i^{th} and j^{th} particles. ε_0 and ε_r are permittivity in vacuum and relativity permittivity respectively. $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ is the magnitude of the distance between i^{th} and j^{h} particles.

A.3 Long-range interactions

Long-range interactions decay as r^{-n}, where n is the dimensionality of the system. It is very important to appropriately model the long-range interactions for charged species as their range is often greater than half of the box length. The charge-charge interaction, decaying as $\frac{1}{r}$ is very difficult to handle. Different methods are developed for handling the long-range interactions such as Ewald summations, Particle mesh Ewald (PME) method, the reaction field method etc. We use PME method for long-range interactions.

In the PME summation, the long-range interaction is split into short range and long-range contributions. The short-range contribution is estimated in real space, whereas the long-range contribution is calculated in Fourier space. PME method uses discrete Fast-Fourier transforms (FFT) to approximate the reciprocal-space term of the standard Ewald summation by a discrete convolution on an interpolating grid.

2.6 References

(1) Aggarwal, A. K. Structure and Function of Restriction Endonucleases. *Curr. Opin. Struct. Biol.* **1995**, *5* (1), 11–19.

(2) Williams, R. J. Restriction Endonuclease. *Mol. Biotechnol.* **2003**, *23* (3), 225–243.

(3) Sistla, S.; Rao, D. N. S-Adenosyl-L-Methionine–Dependent Restriction Enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39* (1), 1–19.

(4) Old, R. W.; Primrose, S. B. *Principles of Gene Manipulation: An Introduction to Genetic Engineering*; Univ of California Press, 1981; Vol. 2.

(5) Bloom, M. V.; Freyer, G. A.; Micklos, D. A. *Laboratory DNA Science: An Introduction to Recombinant DNA Techniques and Methods of Genome Analysis.*; Benjamin/Cummings Publishing Company, Inc., 1996.

(6) Loenen, W. A.; Dryden, D. T.; Raleigh, E. A.; Wilson, G. G. Type I Restriction Enzymes and Their Relatives. *Nucleic Acids Res.* **2013**, *42* (1), 20–44.

(7) Vipond, I. B.; Baldwin, G. S.; Halford, S. E. Divalent Metal Ions at the Active Sites of the EcoRV and EcoRI Restriction Endonucleases. *Biochemistry* **1995**, *34* (2), 697–704.

(8) Pingoud, A.; Jeltsch, A. Structure and Function of Type II Restriction Endonucleases. *Nucleic Acids Res.* **2001**, *29* (18), 3705–3727.

(9) Thielking, V.; Selent, U.; Koehler, E.; Landgraf, A.; Wolfes, H.; Alves, J.; Pingoud, A. Magnesium (2+) Confers DNA Binding Specificity to the EcoRV Restriction Endonuclease. *Biochemistry* **1992**, *31* (15), 3727–3732.

(10) Pingoud, A.; Jeltsch, A. Recognition and Cleavage of DNA by Type-II Restriction Endonucleases. *Eur. J. Biochem.* **1997**, *246* (1), 1–22.

(11) Belkebir, A.; Azeddoug, H. Metal Ion Dependence of DNA Cleavage by SepMI and EhoI Restriction Endonucleases. *Microbiol. Res.* **2013**, *168* (2), 99–105.

(12) Vipond, I. B.; Halford, S. E. Specific DNA Recognition by EcoRV Restriction Endonuclease Induced by Calcium Ions. *Biochemistry* **1995**, *34* (4), 1113–1119.

(13) Lagunavicius, A.; Grazulis, S.; Balciunaite, E.; Vainius, D.; Siksnys, V. DNA Binding Specificity of Mun I Restriction Endonuclease Is Controlled by PH and Calcium Ions: Involvement of Active Site Carboxylate Residues. *Biochemistry* **1997**, *36* (37), 11093–11099.

(14) Zahran, M.; Berezniak, T.; Imhof, P.; Smith, J. C. Role of Magnesium Ions in DNA Recognition by the EcoRV Restriction Endonuclease. *FEBS Lett.* **2011**, *585* (17), 2739–2743.

(15) Onishi, I.; Sunaba, S.; Yoshida, N.; Hirata, F.; Irisa, M. Role of Mg2+ Ions in DNA Hydrolysis by Eco RV, Studied by the 3D-Reference Interaction Site Model and Molecular Dynamics. *J. Phys. Chem. B* **2018**, *122* (39), 9061–9075.

(16) Aboelnga, M. M.; Wetmore, S. D. Unveiling A Single-Metal Mediated Phosphodiester Bond Cleavage Mechanism for Nucleic Acids: A Multiscale Computational Investigation of a Human DNA Repair Enzyme. *J. Am. Chem. Soc.* **2019**.

(17) Pingoud, V.; Wende, W.; Friedhoff, P.; Reuter, M.; Alves, J.; Jeltsch, A.; Mones, L.; Fuxreiter, M.; Pingoud, A. On the Divalent Metal Ion Dependence of DNA Cleavage by Restriction Endonucleases of the EcoRI Family. *J. Mol. Biol.* **2009**, *393* (1), 140–160.

(18) Taylor, J. D.; Badcoe, I. G.; Clarke, A. R.; Halford, S. E. EcoRV Restriction Endonuclease Binds All DNA Sequences with Equal Affinity. *Biochemistry* **1991**, *30* (36), 8743–8753.

(19) Engler, L. E.; Welch, K. K.; Jen-Jacobson, L. Specific Binding by EcoRV Endonuclease to Its DNA Recognition Site GATATC. *J. Mol. Biol.* **1997**, *269* (1), 82–101.

(20) Sidorova, N. Y.; Muradymov, S.; Rau, D. C. Solution Parameters Modulating DNA Binding Specificity of the Restriction Endonuclease EcoRV. *FEBS J.* **2011**, *278* (15), 2713–2727.

(21) Vipond, I. B.; Halford, S. E. Structure–Function Correlation for the EcoRV Restriction Enzyme: From Non-specific Binding to Specific DNA Cleavage. *Mol. Microbiol.* **1993**, *9* (2), 225–231.

(22) Winkler, F. K.; Banner, D. W.; Oefner, C.; Tsernoglou, D.; Brown, R. S.; Heathman, S. P.; Bryan, R. K.; Martin, P. D.; Petratos, K.; Wilson, K. S. The Crystal Structure of EcoRV Endonuclease and of Its Complexes with Cognate and Non-cognate DNA Fragments. *EMBO J.* **1993**, *12* (5), 1781–1795.

(23) Kostrewa, D.; Winkler, F. K. Mg2+ Binding to the Active Site of EcoRV Endonuclease: A Crystallographic Study of Complexes with Substrate and Product DNA at 2-. ANG. Resolution. *Biochemistry* **1995**, *34* (2), 683–696.

(24) Thomas, M. P.; Brady, R. L.; Halford, S. E.; Sessions, R. B.; Baldwin, G. S. Structural Analysis of a Mutational Hot-Spot in the Eco RV Restriction Endonuclease: A Catalytic Role for a Main Chain Carbonyl Group. *Nucleic Acids Res.* **1999**, *27* (17), 3438–3445.

(25) Groll, D. H.; Jeltsch, A.; Selent, U.; Pingoud, A. Does the Restriction Endonuclease Eco RV Employ a Two-Metal-Ion Mechanism for DNA Cleavage? *Biochemistry* **1997**, *36* (38), 11389–11401.

(26) Horton, N. C.; Otey, C.; Lusetti, S.; Sam, M. D.; Kohn, J.; Martin, A. M.; Ananthnarayan, V.; Perona, J. J. Electrostatic Contributions to Site Specific DNA Cleavage by Eco RV Endonuclease. *Biochemistry* **2002**, *41* (35), 10754–10763.

(27) Selent, U.; Rueter, T.; Koehler, E.; Liedtke, M.; Thielking, V.; Alves, J.; Oelgeschlaeger, T.; Wolfes, H.; Peters, F.; Pingoud, A. A Site-Directed Mutagenesis Study to Identify Amino Acid Residues Involved in the Catalytic Function of the Restriction Endonuclease EcoRV. *Biochemistry* **1992**, *31* (20), 4808–4815.

(28) Pingoud, A.; Wilson, G. G.; Wende, W. Type II Restriction Endonucleases—a Historical Perspective and More. *Nucleic Acids Res.* **2014**, *42* (12), 7489–7527.

(29) Jeltsch, A.; Alves, J.; Maass, G.; Pingoud, A. On the Catalytic Mechanism of EcoRI and EcoRV A Detailed Proposal Based on Biochemical Results, Structural Data and Molecular Modelling. *FEBS Lett.* **1992**, *304* (1), 4–8.

(30) Sanchez, D. S.; Zocchi, G. EcoRV Catalysis with a Pre-Bent Substrate. *AIP Adv.* **2015**, *5* (5), 057153.

(31) Sikdar, S.; Ghosh, M.; Adak, A.; Chakrabarti, J. Structural and Dynamic Responses of Calcium Ion Binding Loop Residues in Metallo-Proteins. *Biophys. Chem.* **2019**, 106207.

(32) Mura, C.; McAnany, C. E. An Introduction to Biomolecular Simulations and Docking. *Mol. Simul.* **2014**, *40* (10–11), 732–764.

(33) Sikdar, S.; Chakrabarti, J.; Ghosh, M. A Microscopic Insight from Conformational Thermodynamics to Functional Ligand Binding in Proteins. *Mol. Biosyst.* **2014**, *10* (12), 3280–3289.

(34) Das, A.; Chakrabarti, J.; Ghosh, M. Conformational Contribution to Thermodynamics of Binding in Protein-Peptide Complexes through Microscopic Simulation. *Biophys. J.* **2013**, *104* (6), 1274–1284.

(35) Uzunov, D. I. *Introduction to the Theory of Critical Phenomena: Mean Field, Fluctuations and Renormalization*; World Scientific, 1993.

(36) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational Entropy in Molecular Recognition by Proteins. *Nature* **2007**, *448* (7151), 325.

(37) Dutta, S.; Ghosh, M.; Chakrabarti, J. In-Silico Studies on Conformational Stability of Flagellin–Receptor Complexes. *J. Biomol. Struct. Dyn.* **2019**, 1–13.

(38) Dutta, S.; Ghosh, M.; Chakrabarti, J. Spatio-Temporal Coordination among Functional Residues in Protein. *Sci. Rep.* **2017**, *7*, 40439.

(39) Maganti, L.; Ghosh, M.; Chakrabarti, J. Molecular Dynamics Studies on Conformational Thermodynamics of Orai1–Calmodulin Complex. *J. Biomol. Struct. Dyn.* **2018**, *36* (13), 3411–3419.

(40) Mondal, M.; Chakrabarti, J.; Ghosh, M. Molecular Dynamics Simulations on Interaction between Bacterial Proteins: Implication on Pathogenic Activities. *Proteins Struct. Funct. Bioinforma.* **2018**, *86* (3), 370–378.

(41) Mondal, M.; Halder, S.; Chakrabarti, J.; Bhattacharyya, D. Hybrid Simulation Approach Incorporating Microscopic Interaction along with Rigid Body Degrees of Freedom for Stacking between Base Pairs. *Biopolymers* **2016**, *105* (4), 212–226.

(42) Watson, J. D.; Crick, F. H. Molecular Structure of Nucleic Acids. *Nature* **1953**, *171* (4356), 737–738.

(43) Chandrasekaran, R.; Arnott, S. The Structure of B-DNA in Oriented Fibers. *J. Biomol. Struct. Dyn.* **1996**, *13* (6), 1015–1027.

(44) Bansal, M.; Bhattacharyya, D.; Ravi, B. NUPARM and NUCGEN: Software for Analysis and Generation of Sequence Dependent Nucleic Acid Structures. *Bioinformatics* **1995**, *11* (3), 281–287.

(45) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.

(46) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham III, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, *92* (11), 3817–3829.

(47) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W. AMBER 2015; University of California: San Francisco, CA, 2015. *There No Corresp. Rec. This Ref. Sch.* **2015**.

(48) Debye, P. J. W. Collected Papers of Peter JW Debye. **1954**.

(49) Hestenes, M. R.; Stiefel, E. *Methods of Conjugate Gradients for Solving Linear Systems*; NBS Washington, DC, 1952; Vol. 49.

(50) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(51) Biovia, D. S. *Discovery Studio Modeling Environment*; Release, 2017.

(52) Ho, P. S.; Carter, M. DNA Structure: Alphabet Soup for the Cellular Soul. In *DNA Replication-Current Advances*; IntechOpen, 2011.

(53) Pearson, K. *The Life, Letters and Labours of Francis Galton*; Cambridge University Press, 2011.

(54) Kendall, M. G. Rank Correlation Methods. **1948**.

(55) Jeltsch, A.; Alves, J.; Wolfes, H.; Maass, G.; Pingoud, A. Substrate-Assisted Catalysis in the Cleavage of DNA by the EcoRI and EcoRV Restriction Enzymes. *Proc. Natl. Acad. Sci.* **1993**, *90* (18), 8499–8503.

(56) Anderson, J. E. Restriction Endonucleases and Modification Methylases. *Curr. Opin. Struct. Biol.* **1993**, *3* (1), 24–30.

(57) Thielking, V.; Selent, U.; Koehler, E.; Wolfes, H.; Pieper, U.; Geiger, R.; Urbanke, C.; Winkler, F. K.; Pingoud, A. Site-Directed Mutagenesis Studies with EcoRV Restriction Endonuclease to Identify Regions Involved in Recognition and Catalysis. *Biochemistry* **1991**, *30* (26), 6416–6422.

(58) Loenen, W. A.; Dryden, D. T.; Raleigh, E. A.; Wilson, G. G.; Murray, N. E. Highlights of the DNA Cutters: A Short History of the Restriction Enzymes. *Nucleic Acids Res.* **2013**, *42* (1), 3–19.

(59) Kovall, R. A.; Matthews, B. W. Structural, Functional, and Evolutionary Relationships between λ-Exonuclease and the Type II Restriction Endonucleases. *Proc. Natl. Acad. Sci.* **1998**, *95* (14), 7893–7897.

(60) Ban, C.; Yang, W. Structural Basis for MutH Activation in E. Coli Mismatch Repair and Relationship of MutH to Restriction Endonucleases. *EMBO J.* **1998**, *17* (5), 1526–1534.

CHAPTER 3

Cofactor and Sequence Sensitivity in DNA Cleavage by EcoRI

3.1 Introduction

Here we study another type II RE, namely, EcoRI. EcoRI is a dimeric protein, having two identical chains A and B. Each chain consists of 261 residues. The recognition sequence of EcoRI is GAATTC, and the cleavage occurs between G and A on both strands in presence of Mg^{2+} ions.[1–5] The residues ASP75 and GLU95 in EcoRI hold the metal ion.[2,6] Several factors are important to the cleavage as detailed below:

1. Experimental and theoretical studies suggest that only Mg^{2+}acts as a co-factor as in EcoRV detailed in the previous chapter. Mg^{2+} might have several functions in EcoRI: It might aid in the binding and positioning of DNA, as well as partially neutralizing the extra negative charge in the transition state and protonate the leaving group after cleavage by providing a water molecule in its coordination sphere.[6]
2. Mutation of the DNA sequence also affects the cleavage activity. When cognate sequence (GAATTC) of EcoRI differ by a single base pair from this sequence, the site is termed as EcoRI* site. EcoRI* sites appear in natural DNA more frequently than the canonical sites. Experimental study shows that EcoRI is able to discriminate very accurately between its cognate sequence and the sequence differing by one base pair (EcoRI* site) under normal buffer conditions.[7]
3. The role of water molecules in DNA cleavage by RE has been emphasized in experiment. The cleavage of DNA by EcoRI involves an attack of a water molecule on the scissile phosphate group. Experimental study shows that the pro-Rp phosphoryl oxygen of the phosphate group on the 3' side of the scissile bond activates the attacking water molecule.[6] EcoRI binding to the cognate sequence (GAATTC) releases 70 fewer

water molecules than binding to a noncognate sequence (TAATTC), according to another experiment.[8] Furthermore, the cleavage reaction is accompanied by the binding of 30 water molecules in cognate sequence.[8] Molecular dynamics simulations study proposes that one of the crystal water coordinated with Mg^{2+} attacks the scissile phosphate, making Mg^{2+} to behave as a Lewis acid. However, it is yet to be understood why the water molecule coordinated with Ca^{2+} in Ca^{2+} loaded EcoRI-DNA is unable to attack the scissile phosphate group like Mg^{2+} does in Mg^{2+} loaded EcoRI-DNA in order to cleave the DNA.

We report in Chapter 2 molecular dynamics (MD) simulations studies which show that replacement of Mg^{2+} ions by Ca^{2+} ions result in the cleavage region in EcoRV-DNA complex becoming unstable and disordered due to enhanced fluctuations of one of the catalytically active residues and the cleavage base pairs of the DNA.[9] In this chapter our primary aim is to describe the sensitivity to metal ion cofactor and DNA sequence with emphasis on hydration level. We consider several cases: (1) EcoRI-DNA, (2) EcoRI and cognate sequence in presence of Mg^{2+} ions (Mg^{2+}-EcoRI-DNA), (3) EcoRI and cognate sequence in presence of Ca^{2+} ions (Ca^{2+}-EcoRI-DNA) and (4) Mg^{2+} loaded EcoRI and mutated DNA sequence (Mg^{2+}-EcoRI-mtDNA). We also consider the free states of the components. We compute the conformational thermodynamics of the dihedral angles of EcoRI and the base pair step parameters of DNA in the complex compared to the free DNA, as detailed in chapter 2. We also perform steered molecular dynamics (SMD) simulations and stopped SMD simulations using the SMD snapshots to investigate the dynamic role of water in the DNA recognition region. The conformational thermodynamics data shows that Mg^{2+}-EcoRI-DNA complex has maximum stabilization and order among all complexes. We observe that the number of hydrogen bonded water molecules around the scissile phosphate is the least in Mg^{2+}-EcoRI-DNA.

3.2 Materials and methods

3.2.1 System preparations

The initial structure of free DNA that contains the EcoRI recognition sequence (GAATTC) is modelled by using nucgen[10] software. The DNA bases sequence is 5'-dCGC**GAATTC**GCG-3', with the recognition sequence shown in bold, in 5'-3' direction and the complementary sequence in 3'-5' direction. The initial structure of EcoRI is obtained from protein data bank (PDB) (PDB ID: 1QC9). The metal ions, such as Mg^{2+} and Ca^{2+} loaded EcoRI (Mg^{2+}-EcoRI and Ca^{2+}-EcoRI) are not available in PDB. So, we model Mg^{2+}-EcoRI complex by inserting one Mg^{2+} ion in the vicinity of ASP75 and GLU95 residues as these residues are reported to be the catalytically active residues, in each subunit of EcoRI. We model Ca^{2+}-EcoRI complex the same way as we model Mg^{2+}-EcoRI complex. The crystal structure of EcoRI bound with DNA (EcoRI-DNA) is also taken from protein data bank (PDB) with PDB ID 1ERI[11]. In this complex, the sequence of the DNA bases is same to that of free DNA in both strands.

The complexes of metal ions, such as Mg^{2+} and Ca^{2+} loaded EcoRI-DNA are not available. We model Mg^{2+}-EcoRI-DNA complex by inserting two Mg^{2+} ions in the vicinity of the two active sites of EcoRI-DNA. Similarly, we model Ca^{2+}-EcoRI-DNA complex by placing two Ca^{2+} ions in the vicinity of the two active sites of EcoRI-DNA. We also model Mg^{2+}-EcoRI-mtDNA complex, where we mutate the first A:T base pair in recognition sequence by G:C base pair in Mg^{2+}-EcoRI-DNA complex. As the two active sites are symmetrical in all the complexes, we consider in our study the first active site involving chain A of EcoRI.

3.2.2 Molecular dynamics (MD) simulations

We follow similar protocol as detailed in the Materials and methods of Chapter 2 to perform MD simulations.

3.2.3 Steered molecular dynamics (SMD) simulations

We carry out SMD[12] simulations by applying a force to one end of the DNA along the DNA helix, which is aligned along z axis. The α carbons of EcoRI residues are restrained by a harmonic potential of force constant, k_{res}=10 kcal mol^{-1} nm^{-2}, while the protein is able to move freely along x and y directions. The SMD force that varies with time, is calculated by the following formula, $F(t) = k_{SMD}[z_{SMD}(t) - z_{COM}(t)]$, where $z_{COM}(t)$ is the z coordinate of the center of mass (COM) of DNA, $z_{SMD}(t) = z_{COM}(0) + v_{SMD}t$. v_{SMD} = 0.5 nm ns^{-1} is the SMD velocity, and k_{SMD}= 50,000 kcal mol^{-1} nm^{-2} is the force constant.

In addition to the SMD simulations, we carry out a series of stopped SMD simulations. A stopped SMD simulation is the continuation of a constant-velocity SMD with the velocity set to zero. A stopped SMD simulation is described by the reference SMD simulation time t where the complex is subjected to a high strain rate. We carry out both SMD and stopped SMD simulations on EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes.

3.2.4 Calculations of DNA base pair step parameters and conformational thermodynamics

The detailed calculations of the DNA base pair step parameters and the calculations of the changes in conformational thermodynamics is discussed in the Materials and methods section in Chapter 2.

3.2.5 Calculations of density profile

The density profile is calculated by the following formula,

$$\rho(r) = \frac{n(r)m}{V}$$

We calculate $n(r)$ the number of water molecules within r and $r + dr$ around the metal ion, we consider a shell of thickness dr around the metal ion. Here *V* is the volume of the shell and m is the molecular mass. Thus,

$$\rho(r) = \frac{n(r)m}{[\left(\frac{4}{3}\right)\pi(r+dr)^3 - \left(\frac{4}{3}\right)\pi r^3]}$$

3.2.6 Displacements of nucleotides

The displacement of nucleotide from the interaction regime of EcoRI is characterized as follows: first, for each nucleotide, we calculate the distance between each atom of a particular nucleotide and each atom of EcoRI residues. If the distance is less than 0.45 nm, we make a list of the atom pairs (ij), where atom i belongs to EcoRI, and atom j belongs to DNA. We compute the average distance of the atom pairs of each nucleotide in a single snapshot, denoted as d_{ij}. Finally, we calculate the average distance ($\frac{1}{n} < d_{ij} >$) of the atom pairs of each nucleotide over the last 10 ns of the equilibration. n is the number of pairs. We also calculate d_{ij} from the SMD trajectory for each snapshot. Finally, the displacement of each nucleotide is computed as $D = \frac{1}{n}\sum_{ij}(d_{ij} - < d_{ij} >)$. We consider a nucleotide being ruptured if $D > 0.25$ nm during the SMD simulation.

3.3 Results

Fig. 1(a) depicts the numbering of phosphate, sugar and bases of DNA. The recognition sequence is shown by the box in dotted line. There are two active sites in EcoRI-DNA complex, one containing G_4:C_{21} and A_5:T_{20} base pairs, and another one containing T_8:A_{17} and C_9:G_{16} base pairs. The arrows in the figure shows the cleavage points. Both ASP75 and GLU95 in chain A for the first active site and the same residues in chain B for the second active site are known to be the catalytically active residues.

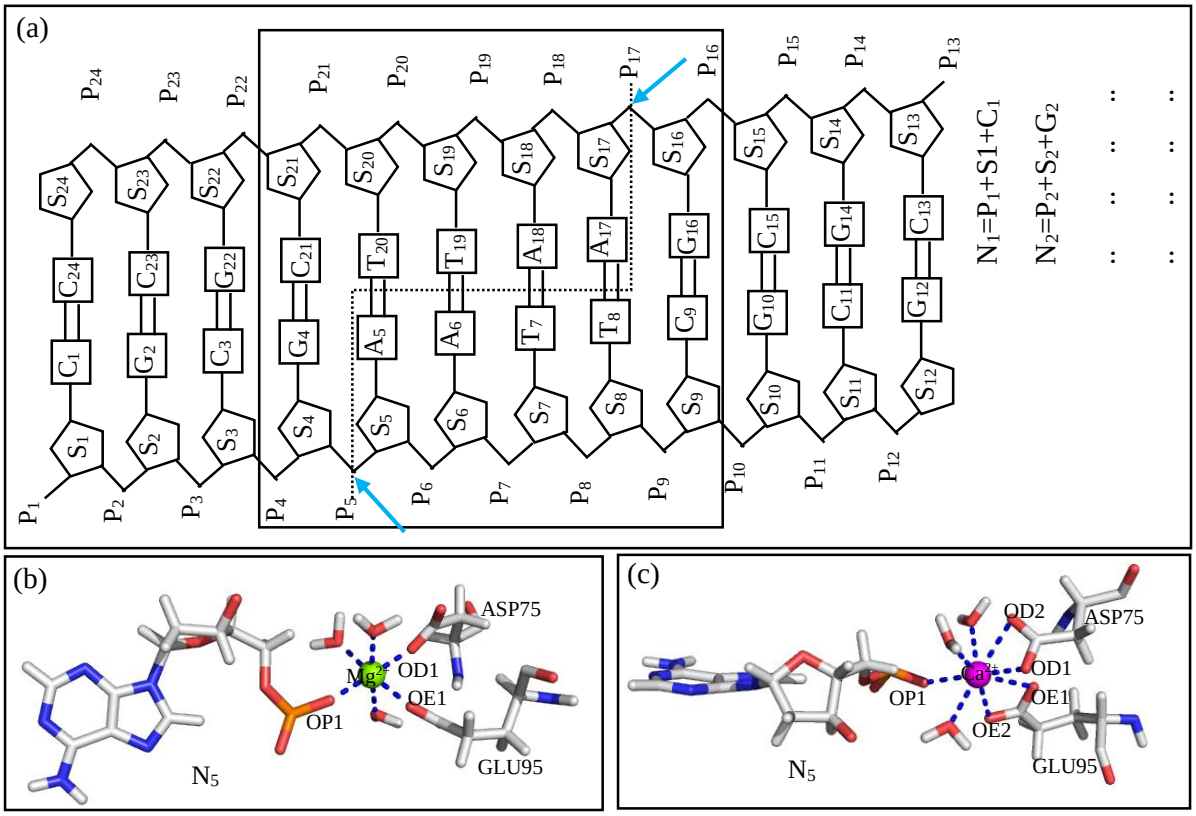

Fig. 1. (a) Labelling of DNA functional groups. N_1, N_2, etc. are the nucleotides where each of them contains a phosphate group, a sugar group and one of the four bases. Coordinations of (b) Mg^{2+} in Mg^{2+}-EcoRI-DNA and (c) Ca^{2+} in Ca^{2+}-EcoRI-DNA.

We carry out MD simulations (as detailed in Materials and methods section of Chapter 2). We show the coordinations of Mg^{2+} in Mg^{2+}-EcoRI-DNA complex, and the coordinations of Ca^{2+} in Ca^{2+}-EcoRI-DNA complex, shown in Figs. 1(b) and (c) respectively. If the distance between any atom of DNA or EcoRI and Mg^{2+} or Ca^{2+} is less or equal to 3.0 Å, coordination is considered. OD1 of ASP75, OE1 of GLU95 and OP1 of P_5 participate in coordination with Mg^{2+} in Mg^{2+}-EcoRI-DNA. Similarly, OD1 and OD2 of ASP75, OE1 and OE2 of GLU95 and OP1 of P_5 participate in coordination with Ca^{2+}. The rest of the coordinations of Mg^{2+} and Ca^{2+} are satisfied by the O atoms of water molecules.

3.3.1 Microscopic fluctuations of conformational variables

We calculate fluctuations of conformational variables, like base pair step parameters of DNA base pair steps and the protein dihedrals as detailed in chapter 2. Here, we show data for the fluctuations of base pair step, (G_4:C_{21}, A_5:T_{20}) and the EcoRI residues, ASP75 and GLU95, belonging to the cleavage region of the first active site. For a given base pair step (i and i+1), the distribution is denoted by $H^{f}_{i,i+1}(b)$ in the free state and $H^{c}_{i,i+1}(b)$ in the complex for a given base pair step parameter *b* introduced in the previous chapter. Here c = MF for EcoRI-DNA, c = Mg for Mg^{2+}-EcoRI-DNA, c = Ca for Ca^{2+}-EcoRI-DNA and c = Mut for Mg^{2+}-EcoRI-mtDNA. Similarly, for a given dihedral ξ of a residue R, the distribution in free state is $H^{(f)}_{R}(\xi)$ and in complex, $H^{(c)}_{R}(\xi)$.

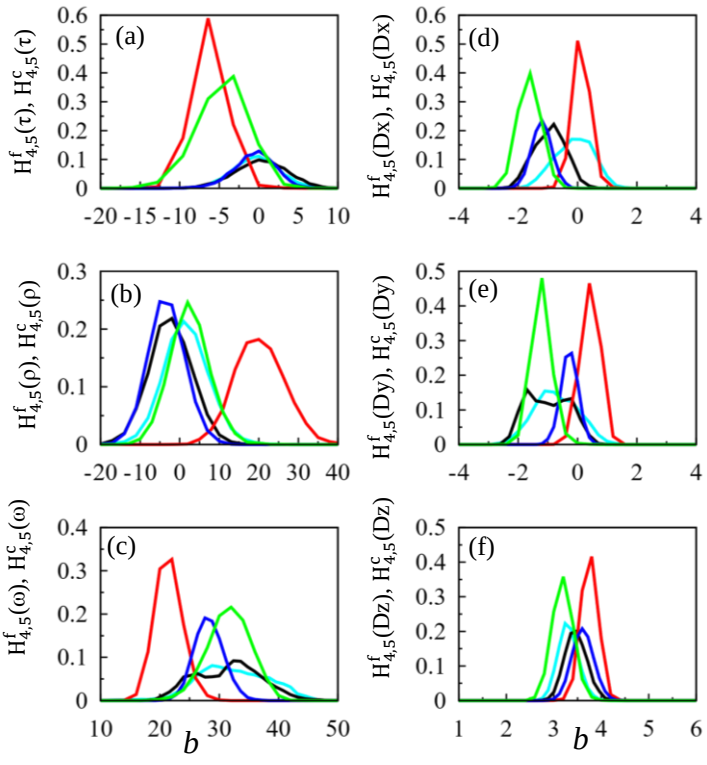

Fig. 2. Distributions of base pair step parameters of DNA in free state (cyan), EcoRI-DNA (black), Mg^{2+}-EcoRI-DNA (red), Ca^{2+}-EcoRI-DNA (blue) and Mg^{2+}-EcoRI-mtDNA (green). The histograms are: (a) τ, (b) ρ, (c) ω, (d) Dx, (e) Dy and (f) Dz of (4,5) base pair step.

We consider the distributions of DNA base pair step parameters. The distributions of τ are single peaked for Mg^{2+} -EcoRI-DNA and for Mg^{2+} -EcoRI-mtDNA, but broad for the

other two complexes (Fig. 2 (a)). The other rotational degree of freedom ρ show single peaks in free state as well as in all the complexes. So far as the twist is concerned, $H^{f}_{4,5}(\omega)$ and $H^{MF}_{4,5}(\omega)$ are multimodal distributions, while $H^{Mg}_{4,5}(\omega)$, $H^{Mut}_{4,5}(\omega)$ and $H^{Ca}_{4,5}(\omega)$ show single peaked distributions (Fig. 2(c)). Both $H^{f}_{4,5}$(Dx), $H^{MF}_{4,5}$(Dx) and $H^{Ca}_{4,5}$(Dx) show broad distributions, while $H^{Mg}_{4,5}$(Dx) and $H^{Mut}_{4,5}$(Dx) are sharp single peaked (Fig. 2(d)). $H^{f}_{4,5}$(Dy) is broad single peaked (Fig. 2(e)). $H^{MF}_{4,5}$(Dy) shows multimodal distribution (Fig. 2(e)). $H^{Mg}_{4,5}$(Dy), $H^{Mut}_{4,5}$(Dy) and $H^{Ca}_{4,5}$(Dy) are all single peaked (Fig. 2(e)). The distribution of Dz shows single peaked distributions for all the cases (Fig. 2(f)). Overall, we observe that the distributions of rotational parameters of the base pair steps are broader in all cases than those of translational parameters.

The distributions of the dihedral angles of ASP75 in free EcoRI and EcoRI in the complexes are shown in Figs. 3(a)-(d). The backbone dihedral φ of ASP75 shows single peaked distributions for all cases (Fig. 3(a)), while the other backbone dihedral ψ shows bimodal distribution in free state and single peaked distributions in the complexes (Fig. 3(b)). $H^{(MF)}_{ASP75}(\chi_1)$ of ASP75 in EcoRI-DNA is single peaked, while $H^{(f)}_{ASP75}(\chi_1)$ shows multimodal distribution (Fig. 3(c)). $H^{(Mg)}_{ASP75}(\chi_1)$ $H^{(Ca)}_{ASP75}(\chi_1)$ $H^{(Mt)}_{ASP75}(\chi_1)$ are all multimodal distributions (Fig. 3(c)). $H^{(f)}_{ASP75}(\chi_2)$ shows multimodal distribution (Fig. 3(d)). $H^{(MF)}_{ASP75}(\chi_2)$ and $H^{(Mg)}_{ASP75}(\chi_2)$ are bimodal distributions, while $H^{(Ca)}_{ASP75}(\chi_2)$ and $H^{(Mt)}_{ASP75}(\chi_2)$ show multimodal distributions. The distributions of the dihedral angles for GLU95 are shown in Figs. 3(e)-(h). φ of GLU95 shows single peaked distributions for all cases (Fig. 3(e)). However, $H^{(Ca)}_{GLU95}(\varphi)$ shows smaller peak height with broad peak compared to other cases. Similarly, $H^{(f)}_{GLU95}(\psi)$, $H^{(MF)}_{GLU95}(\psi)$, $H^{(Mg)}_{GLU95}(\psi)$, $H^{(Ca)}_{GLU95}(\psi)$ and $H^{(Mt)}_{GLU95}(\psi)$ (Fig. 3(f)) are all single peaked with similar peak heights. $H^{(f)}_{GLU95}(\chi_1)$ shows multimodal distribution (Fig. 3(g)), while $H^{(MF)}_{GLU95}(\chi_1)$, $H^{(Ca)}_{GLU95}(\chi_1)$ and $H^{(Mt)}_{GLU95}(\chi_1)$ are single peaked (Fig. 3(g)). $H^{(Mg)}_{GLU95}(\chi_1)$ shows multimodal distribution.

$H^{(f)}_{GLU95}(\chi_2)$ ((Fig. 3(h))) is multimodal distribution. $H^{(MF)}_{GLU95}(\chi_2)$, $H^{(Mg)}_{GLU95}(\chi_2)$, $H^{(Ca)}_{GLU95}(\chi_2)$ and $H^{(Mt)}_{GLU95}(\chi_2)$ are all single peaked distributions (Fig. 3(h)). All the backbone dihedral angles of the residues show single peaked distributions except $H^{(f)}_{ASP75}(\psi)$ which shows bimodal distribution. The side chain dihedrals of ASP75 for most of the cases show multimodal distributions, while side chain dihedrals of GLU95 for most cases show single peaked distributions.

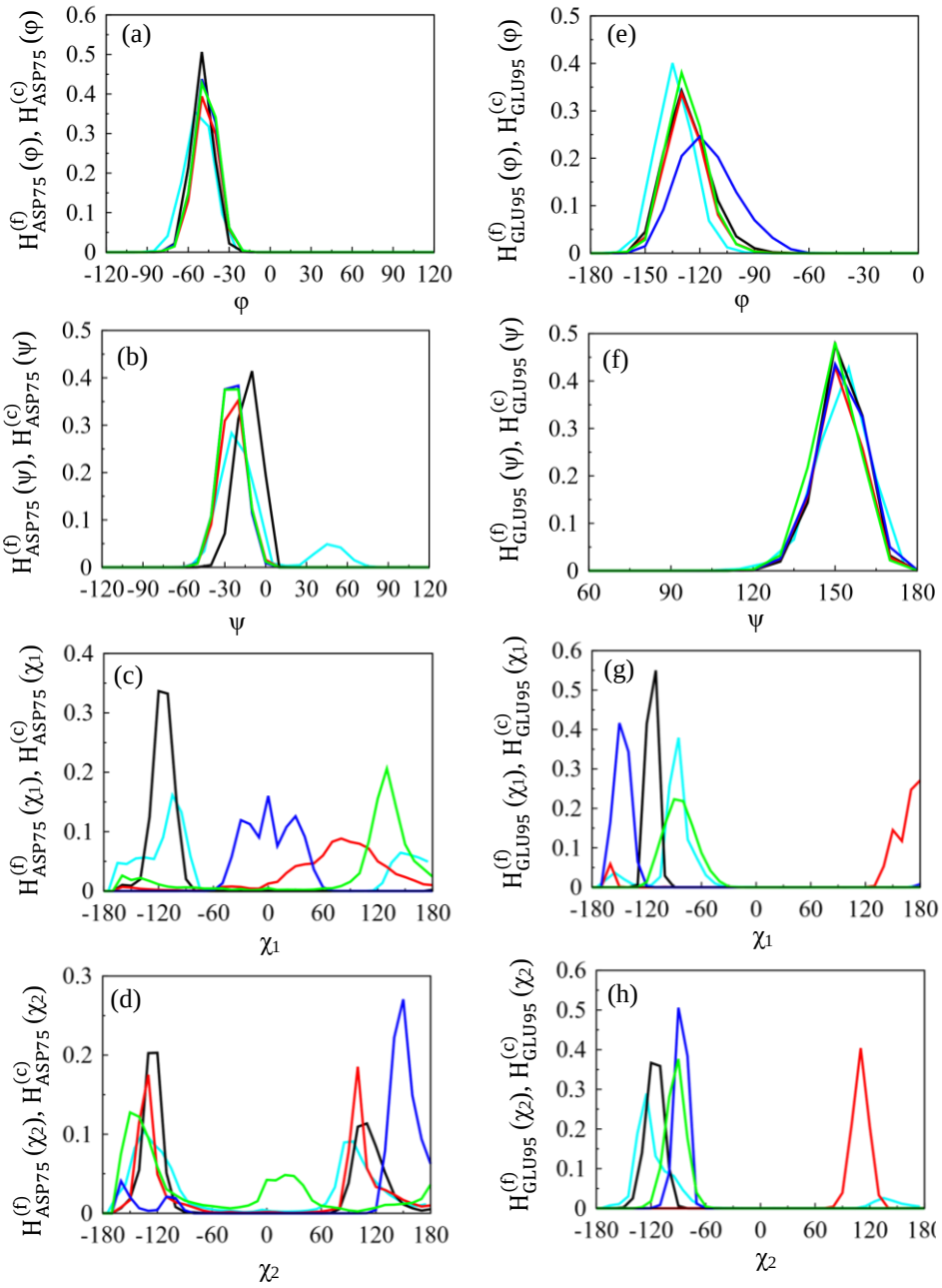

Fig. 3. Histograms of the dihedral angles of EcoRI in free case (cyan), EcoRI-DNA (black), Mg^{2+}-EcoRI-DNA (red), Ca^{2+}-EcoRI-DNA (blue) and Mg^{2+}-EcoRI-mtDNA (green). Distributions of (a) φ, (b) ψ, (c) χ_1 and (d) χ_2 of ASP75. Histograms of (e) φ, (f) ψ, (g) χ_1 and (h) χ_2 of GLU95.

3.3.2 Metal ion selectivity

We compute the changes in the conformational free energy and entropy (as detailed in the Materials and methods of Chapter 2) in the complex with respect to the free state. We

compute the statistical correlation between all variables.[13] We observe that the correlation between the variables is small (correlation coefficient less than 0.5). Hence, we consider the changes in the components separately as in the previous chapter. Here, we compute the changes in conformational free energy and entropy in the complexes with respect to the free state. For a given base pair step parameter, the changes in free energy and entropy of $(\mathrm{i}, \mathrm{i}+1)$ step in the complex with respect to the free state are denoted as $\Delta \mathrm{G}_{\mathrm{i},\mathrm{i}+1}^{\mathrm{cf}}(b)$ and $\mathrm{T}\Delta \mathrm{S}_{\mathrm{i},\mathrm{i}+1}^{\mathrm{cf}}(b)$ respectively, the change in a given base pair step is given by the sum of the individual contributions.

The total free energy and entropy changes of the DNA base pair step parameters for the complexes with respect to free DNA are shown in Figs. 4(a) and (b) respectively. All the base pair steps in recognition region show marginal changes in both free energy and entropy in EcoRI-DNA complex (Fig. 4(a)). All the base pair steps are highly stabilized and ordered in Mg^{2+}-EcoRI-DNA. However, the base pair step (4,5) in cleavage region shows maximum conformational stabilization and order compared to those of the other base pair steps. In Ca^{2+}-EcoRI-DNA complex, the (6,7) step is destabilized. The other base pair steps in this complex show marginal changes (< 2.5 kJ/mol). The entropy changes show similar ordering at all the base pair steps. For the cleavage base pair step (4,5), the stability and order in Mg^{2+}-EcoRI-DNA is greater than those in any other complex.

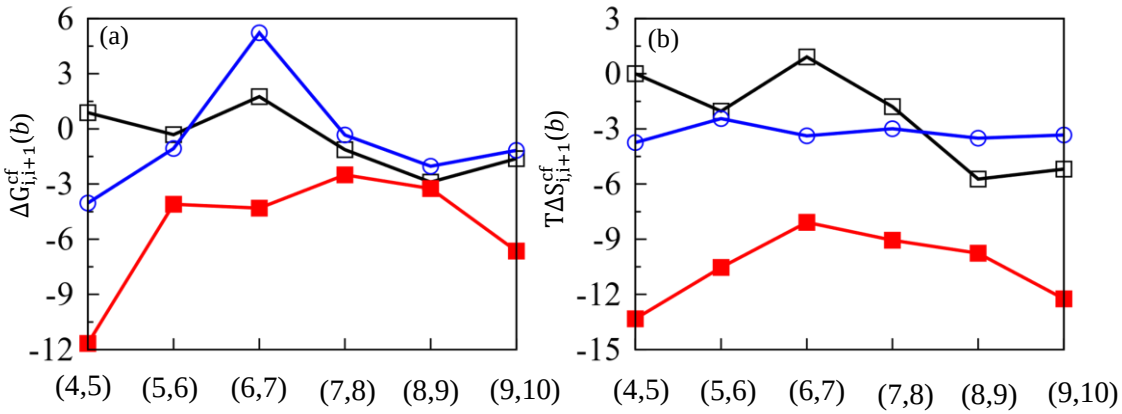

Fig. 4. (a) $\Delta G^{cf}_{i,i+1}$(b) and (b) $T\Delta S^{cf}_{i,i+1}$(b) of base pair step parameters of base pair steps of DNA in EcoRI-DNA (black), Mg^{2+}-EcoRI-DNA (red) and Ca^{2+}-EcoRI-DNA (blue) with respect to free DNA.

We also show the total changes in conformational free energy and entropy for ASP75 and GLU95 residues in EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes with respect to the free EcoRI in Table 1. Both the residues in cleavage region show marginal changes in conformational free energy in all the complexes. However, both the residues are ordered in all the complexes compared to those in free state. The maximum order occurs in GLU95 in Mg^{2+}-EcoRI-DNA complex through χ_1, χ_2 and χ_3. Thus, the overall fluctuations are diminished in the cleavage region in Mg^{2+}-EcoRI-DNA compared to the Ca^{2+}-EcoRI-DNA. This scenario is quite analogous to that of ECORV. The enhanced stability and order is pre-requisite for positioning the metal ion in the vicinity of the scissile phosphate and is more efficiently done in Mg^{2+}-EcoRI-DNA than in Ca^{2+}-EcoRI-DNA.

Table 1. Changes in conformational free energy and entropy of ASP75 and GLU95 in the complexes with respect to free EcoRI.

Residues	System	$\Delta G_R^{cf}(T)$	φ	ψ	χ_1	χ_2	χ_3	$T\Delta S_R^{cf}(T)$	φ	ψ	χ_1	χ_2	χ_3
ASP75	EcoRI-DNA	-2.51	-0.54	-1.43	-0.41	-0.13	0.00	-5.03	-0.54	-1.46	-1.98	-1.06	0.00
	Mg^{2+}-EcoRI-DNA	-1.12	-0.01	-1.02	0.01	-0.09	0.00	-2.79	-0.44	-1.37	0.27	-1.27	0.00
	Ca^{2+}-EcoRI-DNA	-1.76	-0.27	-1.28	-0.08	-0.14	0.00	-4.40	-0.34	-1.49	0.02	-2.59	0.00
	Mg^{2+}-EcoRI-mtDNA	-1.47	-0.25	-1.27	0.02	0.04	0.00	-2.59	-0.37	-1.33	-0.67	-0.22	0.00
GLU95	EcoRI-DNA	-0.55	0.3	-0.24	-0.33	-0.22	-0.06	-3.9	0.46	0.27	-2.14	-1.76	-0.73
	Mg^{2+}-EcoRI-DNA	0.47	0.39	0.03	0.36	-0.22	-0.09	-5.29	0.03	0.33	-1.99	-1.95	-1.70
	Ca^{2+}-EcoRI-DNA	-0.25	0.60	-0.09	-0.19	-0.35	-0.22	-4.07	1.08	0.38	-1.11	-2.65	-1.77
	Mg^{2+}-EcoRI-mtDNA	-0.55	0.09	-0.05	-0.01	-0.21	-0.36	-2.68	0.09	0.44	0.7	-1.43	-2.47

3.3.3 Sequence selectivity

We now focus on how mutation of A:T base pair in recognition sequence by G:C affects the DNA cleavage by EcoRI in presence of Mg^{2+} ions. Here we focus on the conformational thermodynamics changes in the active site protein residues. We consider two catalytically active residues ASP75 and GLU95 in Mg^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes, shown in Table 1. Both the residues show marginal changes in conformational free energy in both complexes. However, the changes in conformational entropy make a difference in these two complexes. ASP75 is ordered through both backbone and side chain dihedrals in both the complexes. The changes in entropy of this residue is similar in both the complexes. However, GLU95 is highly ordered in Mg^{2+}-EcoRI-DNA complex compared to those in Mg^{2+}-EcoRI-mtDNA complex through the side chain dihedrals. So, GLU95 makes the coordination of the

metal ion better in Mg^{2+}-EcoRI-DNA than in the Mg^{2+}-EcoRI-mtDNA complex. This could be the microscopic mechanism behind the sequence selectivity.

3.3.4 Hydration levels

We calculate the water density profiles (details in Materials and methods) around the metal ions in metal ion loaded complexes. Fig. 5(a) shows the distribution of water molecules around Mg^{2+} in both Mg^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes and around Ca^{2+} in Ca^{2+}-EcoRI-DNA complex. For all cases, the distributions show two well defined peaks. The first peak in Mg^{2+}-EcoRI-DNA appears at 0.215 nm, while the second peak arises at 0.285 nm. The density is observed to be maximum of 3.097 gm/cc at 0.215 nm. The distribution in Mg^{2+}-EcoRI-mtDNA is similar as those in Mg^{2+}-EcoRI-DNA. However, the distribution in Ca^{2+}-EcoRI-DNA complex varies widely. As the radius of Ca^{2+} is greater than that of Mg^{2+}, the peaks appear further away from Ca^{2+} than in the case of Mg^{2+}. The first peak arises at 0.25 nm, while the second peak appears at 0.315 nm. Here, the maximum density is 1.982 gm/cc which occurs at 0.25 nm. So, the results suggest that Mg^{2+} is better hydrated than Ca^{2+} due to smaller size and larger charge density.

We also calculate the water distributions around P_5, P_6 and P_7 in EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA near the cleavage region. The density peak around P_5 (Fig. 5(b)) is similar in Mg^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA, and larger compared to the other complexes. Thus, Mg^{2+} ion captures more water molecule due to stronger electrostatic interactions. The density around P_6 (Fig.5(c)) is similar in all the complexes. However, the scenario change for the water distribution around P_7 (Fig.5 (d)). The density of water for Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes shows a significant decrease compared to those for EcoRI-DNA complex.

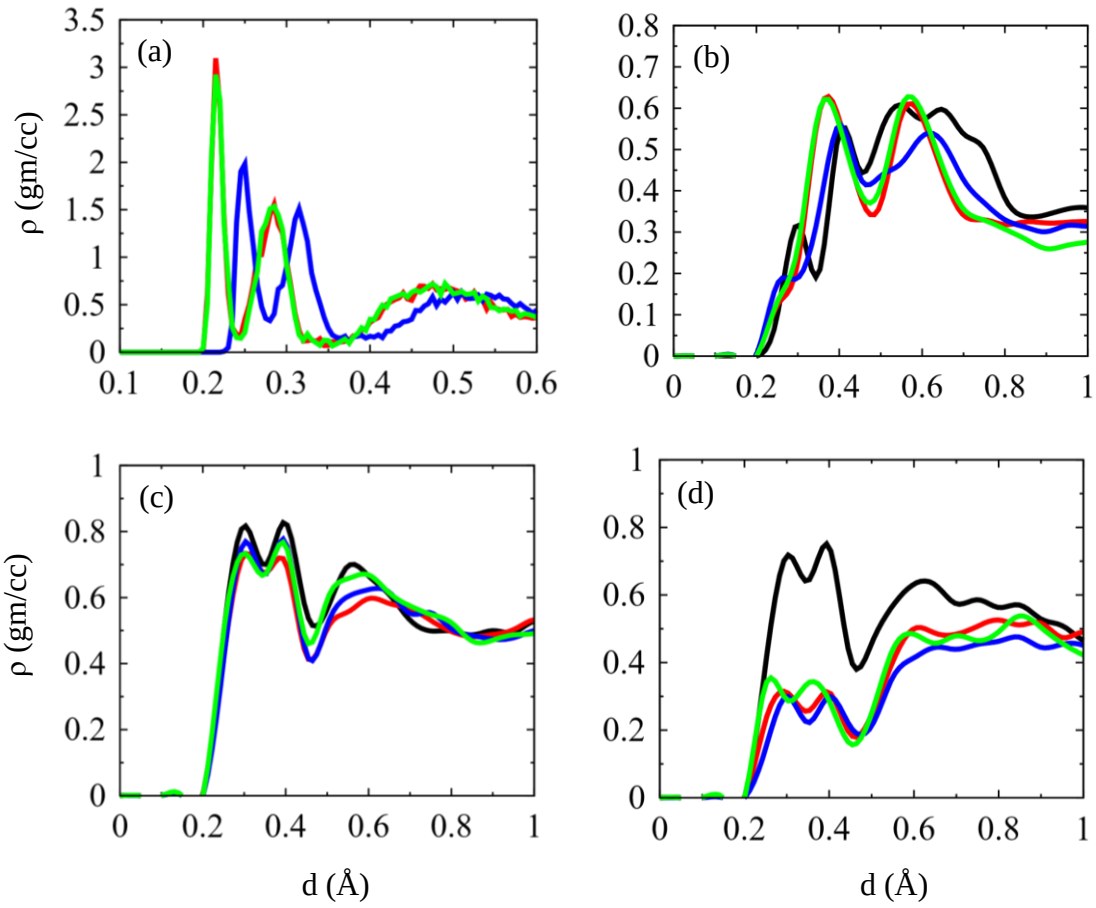

Fig. 5. (a) Density of water, ρ (gm/cc) around Mg^{2+} in Mg^{2+}-EcoRI-DNA (red) and in Mg^{2+}-EcoRI-mtDNA (green), and around Ca^{2+} in Ca^{2+}-EcoRI-DNA (blue). Water density around (b) P_5, (c) P_6 and (d) P_7 for EcoRI-DNA (black), Mg^{2+}-EcoRI-DNA (red), Ca^{2+}-EcoRI-DNA (blue) and Mg^{2+}-EcoRI-mtDNA (green).

We calculate the number of water molecules around each phosphate group, and also calculate number of hydrogen bonds between O atoms of each phosphate group in recognition region and water molecules. For the detection of hydrogen bonds (HBs), we consider the following protocol: (i) the distance between the donor O atom and acceptor O atom is less or equal to 0.34 nm and (ii) O-H (donor) – O (acceptor) angle is less or equal to 150^0. We define *f,* the fraction of hydrogen bonded water molecules out of the available water molecules number. As the two strands of free DNA, EcoRI-DNA, Mg^{2+}-EcoRI-DNA and Ca^{2+}-EcoRI-

DNA complexes are symmetric, the results are averaged over two strands in these complexes, shown in Fig. 6(a). We show results for free DNA as a reference. For free DNA, *f* does not show sequence specificity for different phosphate groups. However, the complexes show sequence specific *f* value. In EcoRI-DNA complex, *f* shows minimum value for the cleavage phosphate group P_5. In Mg^{2+}-EcoRI-DNA complex, the minimum of *f* at P_5 further deeper and a mximum at P_7 unlike EcoRI-DNA complex. In Ca^{2+}-EcoRI-DNA complex, *f* has a minimum at P_5 and a maximum at P_7 but the minimum is comparable with that in EcoRI-DNA complex. As the two strands are asymmetric in Mg^{2+}-EcoRI-mtDNA complex, we separately calculate *f* for each strand, shown in Figs. 6(b) and (c). *f* in both strands show similar profile as those in Mg^{2+}-EcoRI-DNA complex. Thus, the data suggest that the hydrogen bond formation becomes sequence specific in presence of the metal ion.

It may be noted that the number of water molecules surrounding P_5 is more in the presence of Mg^{2+} than the other cases. The minimum of *f* at P_5 in presence of metal ions, indicate that this phosphate group is highly hydrated and more accessible to nucleophilic attack required for cleavage. This is more dominant in presence of Mg^{2+} than Ca^{2+}. The decrease in *f* is compensated by P_7 that belongs to more fluctuating part of the sequence as evident from the conformational thermodynamics data.

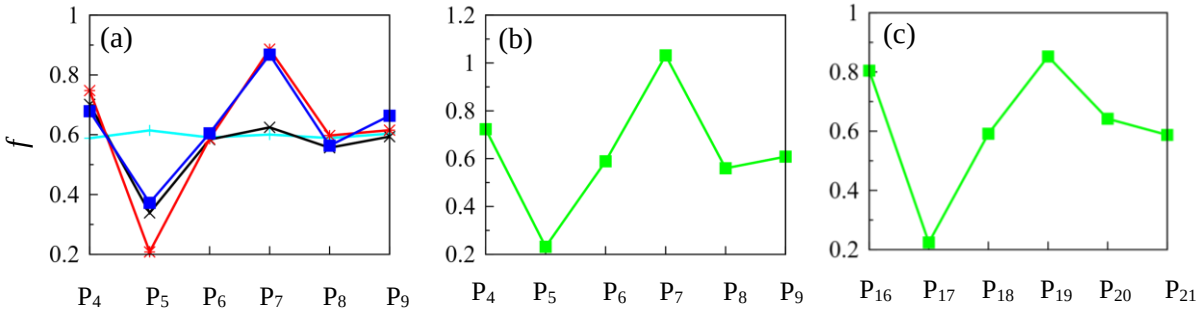

Fig. 6. (a) The number of hydrogen bonds per water molecule, *f* in free DNA (cyan), EcoRI-DNA (black), Mg^{2+}-EcoRI-DNA (red) and Ca^{2+}-EcoRI-DNA (blue). *f* for Mg^{2+}-EcoRI-mtDNA in (b) first strand and (c) second strand.

3.3.5 SMD results

To investigate the effect of unbinding of DNA from the binding pocket of EcoRI, SMD simulations are conducted. We show the force experienced by the system due to unbinding of DNA for EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes in Fig. 7. For EcoRI-DNA, Mg^{2+}-EcoRI-DNA, Ca^{2+}-EcoRI-DNA, the force increases gradually until it reaches a maximum, and after that it gradually decreases. The force is maximum when all the nucleotides of the DNA come out from the interaction regime of protein. The maximum of the force is larger for both Mg^{2+}-EcoRI-DNA than the other case. This indicates that the interactions between EcoRI and DNA is stronger in Mg^{2+}-EcoRI-DNA. The Mg^{2+}-EcoRI-mtDNA complex, on the other hand, shows only a broad maximum which suggests that the binding is rather insensitive compared to the other complexes. The force data indicates that the strongest binding in Mg^{2+}-EcoRI-DNA.

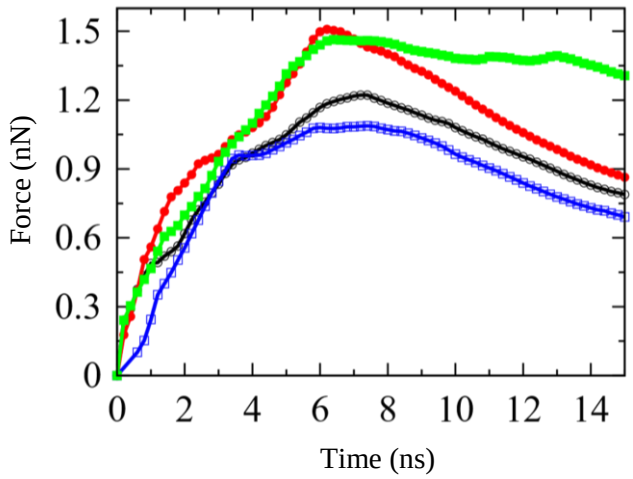

Fig. 7. SMD force profile for EcoRI-DNA (black), Mg^{2+}-EcoRI-DNA (red), Ca^{2+}-EcoRI-DNA (blue) and Mg^{2+}-EcoRI-mtDNA (green).

The displacements of each nucleotide (D) of both strands for different complexes with SMD simulation time are shown in Fig. 8. In the first strand, the nucleotides in the recognition region are N_4-N_9, while the nucleotides in the second strand are N_{16}-N_{21}. The D data of the first strand and the second strand of EcoRI-DNA are shown in Fig. 8(a) and (b) respectively. Nucleotide N_{20} in the second strand is the first nucleotide that breaks away from the interaction

regime of EcoRI around 2 ns. It leaves the binding site before N_{21} which breaks away at ~ 3ns. Nucleotide N_{19} also leaves at same time as N_{21}. After that around 3.8 ns, N_4 in the first strand leaves the binding site. The complex is then stabilized up to 4.7 ns. In the range 5-6 ns, N_5, N_6 and N_7 in the first strand break away in which N_6 and N_7 break away at the same time (~ 5.2 ns), while both N_{18}, N_{17} in the second strand also leave the binding site. The last nucleotide in the recognition region of second strand leaves at 6.6 ns, while that in the first strand breaks away at 6.9 ns. The sequence of nucleotide rupture is similar to the earlier study.[14] However, the rupture time somewhat less in our system which may be due to differences in force fields in two cases.

Next, we consider the two strands (Figs. 8 (c) and (d)) in Mg^{2+}-EcoRI-DNA complex. In the second strand, both the nucleotides N_{21} and N_{20} are ruptured around ~ 3 ns (Fig. 8(d)). N_4 and N_{19} break away at 3.4 ns and 3.9 ns respectively. N_{18} breaks away at 4.95 ns. N_4-N_6 rupture gradually in a gap of ~ 1 ns, while both N_7 and N_8 nucleotides break away approximately at the same time (~ 5.8 ns). The remaining two nucleotides, N_{17} and N_{16} in the second leaves the binding site around ~ 6.5 ns, while the remaining nucleotide N_9 in the first strand breaks away around ~ 8 ns. In EcoRI-DNA, the nucleotide starts breaking away at 2ns, whereas in Mg^{2+}-EcoRI-DNA complex, this time is 3ns. We also observe that the last nucleotide in the recognition region stabilized up to 6.8 ns, while the last nucleotide in Mg^{2+}-EcoRI-DNA complex remains stabilized up to 7.9 ns. It indicates that the interactions between EcoRI and DNA is stronger in Mg^{2+}-EcoRI-DNA than in EcoRI-DNA.

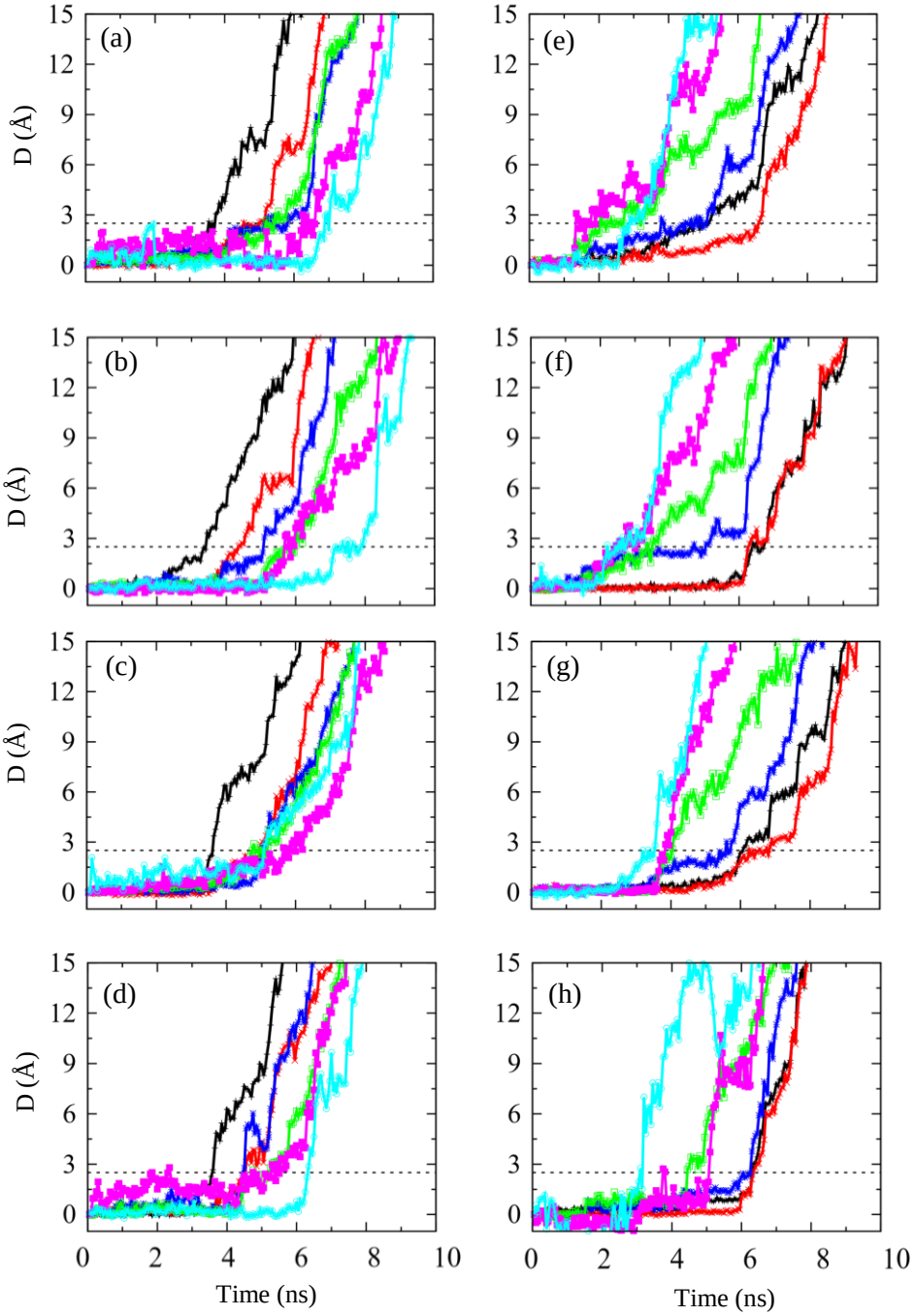

Fig. 8. Displacements (D) of the nucleotides in (a) EcoRI-DNA, (b) Mg^{2+}-EcoRI-DNA, (c) Ca^{2+}-EcoRI-DNA and (d) Mg^{2+}-EcoRI-mtDNA for the first strand. Similarly, the displacements of nucleotides in (e) EcoRI-DNA, (f) Mg^{2+}-EcoRI-DNA, (g) Ca^{2+}-EcoRI-DNA and (h) Mg^{2+}-EcoRI-mtDNA for the second strand. The symbols are as follows: N_4 (black), N_5 (red), N_6 (blue), N_7 (green), N_8 (magenta), N_9 (cyan) in the first strand and N_{16} (black), N_{17} (red), N_{18} (blue), N_{19} (green), N_{20} (magenta) and N_{21} (cyan) in the second strand.

The sequence of events in Ca^{2+}-EcoRI-DNA complex is quite different. D in first and second strands is shown in Fig. 8(e) and (f). In this complex, N_4, N_{20} and N_{21} leaves the binding site approximately at the same time ~ 3.7 ns. N_{19} breaks away at ~ 4.05 ns. After that all four

nucleotides N_5, N_6, N_7 and N_9 in the first strand break away at approximately same time (~ 5.1 ns) (Fig. 8(e)). N_{18} breaks away at ~ 5.6 ns, while N_8 breaks away around ~ 6 ns. Here, N_9 is ruptured before N_8 (Fig. 8(e)). N_{16} comes out from the binding pocket around ~ 6 ns before N_{17} which breaks away at ~ 6.55 ns (Fig. 8(f)). Here, N_{17} breaks away at ~ 6.55 ns, while in Mg^{2+}-EcoRI-DNA the last nucleotide is N_9 that breaks away at ~ 8 ns, indicating Mg^{2+}-EcoRI-DNA complex is more stable than Ca^{2+}-EcoRI-DNA.

In Mg^{2+}-EcoRI-mtDNA complex, in the second strand, the first nucleotide N_{21} (Fig. 8(h) in the recognition region breaks away at ~ 3.2 ns, while N_4 breaks away at ~ 3.6 ns (Fig. 8(g)). Both N_5 and N_6 nucleotides break away at the same time (~ 4.5 ns) (Fig. 8(g)). Around ~ 5 ns, another two nucleotides N_{20} and N_{19} break away. Similarly, both N_7 and N_8 are ruptured at ~ 5.3 ns. The remaining nucleotides N_9, N_{18}, N_{17} and N_{16} break away at the same time (~ 6.35 ns). The results indicate that Mg^{2+}-EcoRI-DNA is more stable than Mg^{2+}-EcoRI-mtDNA.

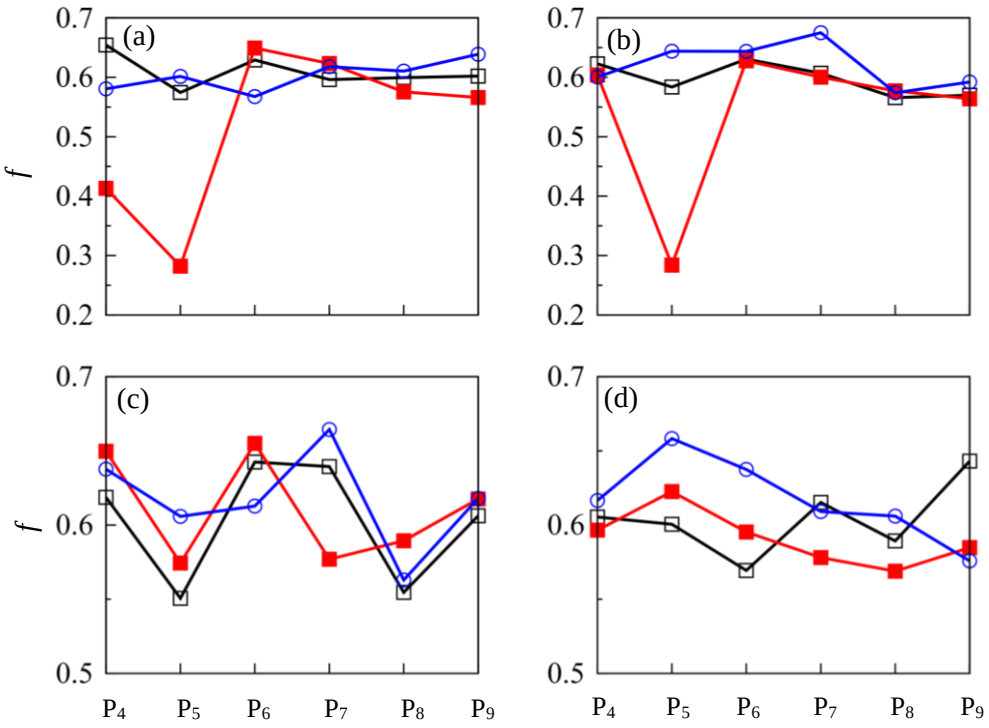

Fig. 9. The number of hydrogen bonds per water molecule, *f* of each phosphate group for stopped SMD simulations at (a) 4 ns, (b) 5 ns, (c) 6ns and (d) 6.5 ns. The colors in black, red and blue illustrate the results for EcoRI-DNA, Mg^{2+}-EcoRI-DNA and Ca^{2+}-EcoRI-DNA respectively.

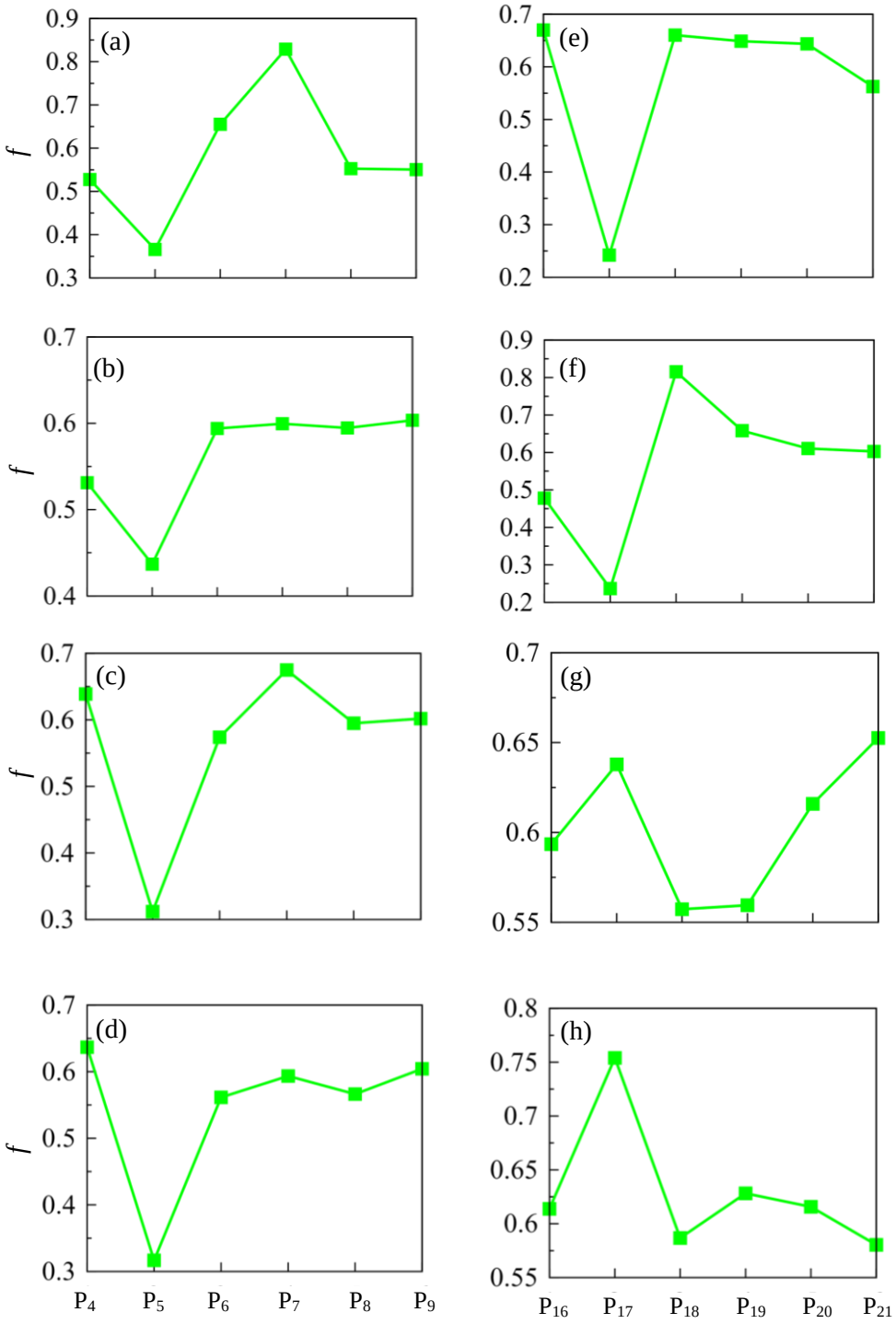

Fig. 10. Number of hydrogen bonds per water molecule, *f* around each phosphate group of Mg^{2+}-EcoRI-mtDNA for stopped SMD simulations at (a) 4 ns, (b) 5ns, (c) 6 ns and (d) 6.5 ns in the first strand. *f* for (e) 4 ns, (f) 5ns, (g) 6 ns and (h) 6.5 ns in the second strand.

We carry out stopped SMD simulations on snapshots from SMD simulation trajectory at 4 ns, 5 ns, 6 ns and 6.5 ns. Each snapshot is used for further simulation of 50 ns and we calculate *f.* At 4 ns (Fig. 9(a)), *f* has similar values for all the phosphate groups in recognition region for EcoRI-DNA and Ca^{2+}-EcoRI-DNA. For Mg^{2+}-EcoRI-DNA complex, *f* has minimum at P_4. Similar behavior is observed at 5 ns of stopped SMD simulations (Fig. 9(b)). At 6 ns and 6.5 ns, the values of *f* get more and more uniform with the SMD time over the phosphates in all the three complexes.

For Mg^{2+}-EcoRI-mtDNA complex, the results are shown for both strands individually in Fig 10 due to asymmetric arrangement of the bases in two strands after the mutation. In the first strand (Figs. 10(a)-(d)) *f* shows minimum at P_5 which does not change much in the first strand. However, in the second strand the minimum at P_5 for 4 ns (Fig.10(e)) and 5 ns (Fig.10(f)) shows tendency to homogenize at 6 ns (Fig. 10(g)) and 6.5 ns (Fig. 10(h)). It is to note that we mutate the second base pair in the recognition region that involves the cleavage region in the first strand, but the cleavage region in the second strand remains same. So, the second strand behaves as that of the cleavage site, while the first strand hardly show any change.

3.4 Conclusions

Our study based on microscopic fluctuations shows that the cleavage region is highly stabilized and ordered both in Mg^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes compared to those in the Ca^{2+}-EcoRI-DNA complex. However, the active protein residues holding the metal ion is less stabilized in Mg^{2+}-EcoRI-mtDNA compared to those in Mg^{2+}-EcoRI-DNA. This suggests that DNA cleavage can take place in the Mg^{2+}-EcoRI-DNA complex. It is also observed that the metal ion mediated hydration of the scissile phosphate group is higher for both Mg^{2+}-EcoRI-DNA and Mg^{2+}-EcoRI-mtDNA complexes compared to those in EcoRI-

DNA and Ca^{2+}-EcoRI-DNA complexes. However, the thermodynamics data show that the cleavage of DNA in Mg^{2+}-EcoRI-mtDNA is less probable than in Mg^{2+}-EcoRI-DNA.

3.5 References

(1) Jeltsch, A.; Alves, J.; Wolfes, H.; Maass, G.; Pingoud, A. Substrate-Assisted Catalysis in the Cleavage of DNA by the EcoRI and EcoRV Restriction Enzymes. *Proc. Natl. Acad. Sci.* **1993**, *90* (18), 8499–8503.

(2) Pingoud, A.; Jeltsch, A. Recognition and Cleavage of DNA by Type-II Restriction Endonucleases. *Eur. J. Biochem.* **1997**, *246* (1), 1–22.

(3) Pingoud, A.; Fuxreiter, M.; Pingoud, V.; Wende, W. Type II Restriction Endonucleases: Structure and Mechanism. *Cell. Mol. Life Sci.* **2005**, *62* (6), 685.

(4) Aggarwal, A. K. Structure and Function of Restriction Endonucleases. *Curr. Opin. Struct. Biol.* **1995**, *5* (1), 11–19.

(5) Anderson, J. E. Restriction Endonucleases and Modification Methylases. *Curr. Opin. Struct. Biol.* **1993**, *3* (1), 24–30.

(6) Jeltsch, A.; Alves, J.; Maass, G.; Pingoud, A. On the Catalytic Mechanism of EcoRI and EcoRV A Detailed Proposal Based on Biochemical Results, Structural Data and Molecular Modelling. *FEBS Lett.* **1992**, *304* (1), 4–8.

(7) Thielking, V.; Alves, J.; Fliess, A.; Maass, G.; Pingoud, A. Accuracy of the EcoRI Restriction Endonuclease: Binding and Cleavage Studies with Oligodeoxynucleotide Substrates Containing Degenerate Recognition Sequences. *Biochemistry* **1990**, *29* (19), 4682–4691.

(8) Robinson, C. R.; Sligar, S. G. Changes in Solvation during DNA Binding and Cleavage Are Critical to Altered Specificity of the EcoRI Endonuclease. *Proc. Natl. Acad. Sci.* **1998**, *95* (5), 2186–2191.

(9) Mandal, S. C.; Maganti, L.; Mondal, M.; Chakrabarti, J. Microscopic Insight to Specificity of Metal Ion Cofactor in DNA Cleavage by Restriction Endonuclease EcoRV. *Biopolymers* e23396.

(10) Bansal, M.; Bhattacharyya, D.; Ravi, B. NUPARM and NUCGEN: Software for Analysis and Generation of Sequence Dependent Nucleic Acid Structures. *Bioinformatics* **1995**, *11* (3), 281–287.

(11) Kim, Y.; Grable, J. C.; Love, R.; Greene, P. J.; Rosenberg, J. M. Refinement of Eco RI Endonuclease Crystal Structure: A Revised Protein Chain Tracing. *Science* **1990**, *249* (4974), 1307–1309.

(12) Israelewitz, B.; Izrailev, S.; Schulten, K. Binding Pathway of Retinal to Bacterio-Opsin: A Prediction by Molecular Dynamics Simulations. *Biophys. J.* **1997**, *73* (6), 2972–2979.

(13) Pearson, K. *The Life, Letters and Labours of Francis Galton*; Cambridge University Press, 2011.

(14) Dorvel, B.; Sigalov, G.; Zhao, Q.; Comer, J.; Dimitrov, V.; Mirsaidov, U.; Aksimentiev, A.; Timp, G. Analyzing the Forces Binding a Restriction Endonuclease to DNA Using a Synthetic Nanopore. *Nucleic Acids Res.* **2009**, *37* (12), 4170–4179.

CHAPTER 4

Quantum Chemical Studies on Interactions of ZnO Nanoparticles with Cellular Energy Carrier Molecules, ATP, ADP and AMP

4.1 Introduction

Nano-bio conjugates stabilized via interactions between inorganic moieties and biomolecules, attract a great deal of attention due to their importance in biomedical applications.[1–5] The bio-conjugate nanomaterials typically use bio-safe metal-oxide nanoparticles.[6] Out of metal-oxide nanoparticles, strongly Raman active ZnO nanoparticles (ZnONP) conjugated with biomolecules,[7–12] like tyrosine/tryptophan,[13] DNA bases,[14] RNA sequences,[10] adenosine triphosphate,[9] are important in various applications including drug carriers, cosmetics and medicinal materials,[15–17] early detection of symptoms of several serious chronic diseases,[13] ultra-fast DNA sequencing devices [18] to name only a few. The conjugation of bio-molecules to nanoparticle surfaces depends on variety of factors, including chemical composition, porosity and surface crystallinity of the nanomaterials.[19–21] Biomolecules may further induce phase transformations, free energy releases, restructuring, and dissolution at the nanomaterial surface.[21] All these factors make microscopic understanding of bio-conjugate materials challenging and lacking, despite intense studies in recent times.[3,22–24]

With high potential biological applications, it is quite imperative to understand how ZnONP interacts with different bio-molecules present in cellular environment. Energy carrier bio-molecules play important role in all cellular metabolic processes. Adenosine triphosphate (ATP), consisting of three phosphate groups, a ribose and an adenine group, participates in biological energy transfer reactions, primarily via controlled hydrolysis of phosphate group forming adenosine diphosphate (ADP) molecule with two phosphate groups where adenosine monophosphate (AMP) with one phosphate group is formed as an intermediate compound.[25]

These bio-molecules have phosphate, adenine and ribose groups. Phosphate groups bind to metal ions by virtue of large valence charge, in particular the divalent metal ions which control of host of processes.[26–30] Raman spectroscopic study shows that Ca^{2+} and Mg^{2+} ions bind strongly with the phosphate group, while Cu^{2+} interacts simultaneously with the phosphate group and the adenine ring of ATP at neutral pH.[31] Chelation has been reported by recent experimental Raman spectroscopic studies of ZnONP and ATP bio-conjugate [9] where in addition to the phosphate groups, adenine moiety also participate in the chelation by the interaction of Zn with nitrogen of adenine.

There is as yet no theoretical study for microscopic understanding of chelation of metal oxide nano particles by the energy carrier bio-molecules. The interactions of biomolecules with nanoparticle surfaces at atomistic level are investigated by quantum chemical calculations using the density functional theory (DFT). The DFT method with B3LYP functional [32] has already been successfully employed to study ground state properties, reaction mechanism, vibrational absorption, and circular dichroism of transition metal complexes having organic ligands, metal-oxide nanoparticles as well as bio molecules [33–35]. It also determines the electrostatic properties which are not readily accessible from experiments, as well as giving additional information about chemical bonding and chemical reactions. As a result such technique has emerged as a good compromise between computational cost, coverage, and accuracy of results.[36]

Here, we carry out DFT based quantum chemical calculations to investigate the nano-bio conjugate consisting of ZnONP with ATP, ADP and AMP. Our objective is to understand microscopic details of chelation by individual groups (phosphate, ribose and adenine respectively) of the ligands with ZnONP. We restrict to small ZnONP, namely, $Zn_{12}O_{12}$ cluster to keep the computation cost manageable. The $Zn_{12}O_{12}$ cluster is stable due to its cage like structure, and have been extensively investigated theoretically and experimentally.[37–41] We not

only check the chelation process reported in the experiments, but also explore its generality across these phosphate containing moieties. In order to emphasize roles of adenine and ribose groups in chelation process, we also consider a model calculation of methyl phosphate (MP) conjugated to the ZnONP which lacks these groups.

4.2 Computational details

All quantum chemical calculations are performed using B3LYP functional[32] within the density functional theory (DFT) implemented in the Gaussian 16 package.[42] The theoretical background of the calculations has been given in Appendix I. The optimizations are done using double zeta quality 6-31G (d) basis set function for the C, N, O and P atoms and 6-31G basis set for the H atoms, while the LANL2DZ[43,44] basis set with effective core potential (ECP)[45] is used for the Zn atoms. For single point energy calculations, triple zeta split valence basis set along with the diffuse and polarization functions, 6-311++G (d,p) is used for the C, N, O, P, and H atoms, while Stuttgart-Dresden basis set with ECP[45] is used for the Zn atoms.

The population analysis is carried out by natural bond orbital (NBO) method[46,47] at B3LYP/6-311++G (d,p) level of theory using NBO program[48] under the Gaussian 16 package.[42] The optimized structural parameters are used to calculate vibrational frequencies to ascertain that each equilibrium geometry does not have any imaginary frequency. These vibrational frequency calculations are done using the same type of basis sets as of geometry optimization.

Binding energy is defined as the differences in total energies of the complex and isolated systems at equilibrium configurations. So, the binding energy of the complex system is determined through the following equation: $E_B = E_{biomolecule-ZnONP} - (E_{biomolecule} + E_{ZnONP})$, where $E_{biomolecule-ZnONP}$ is the total energy of the complex of ZnONP and

biomolecule, E_{ZnONP} and $E_{biomolecule}$ are the total energies of the isolated ZnONP and biomolecule respectively.

4.3 Results

4.3.1 Bonding nature in ground state

The optimized cage structure of $Zn_{12}O_{12}$ cluster, as shown in Fig. 1(a), with six $(ZnO)_2$ and eight $(ZnO)_3$ rings forming a truncated octahedron. Here all Zn and O vertices are equivalent. In $Zn_{12}O_{12}$ cluster, shorter Zn-O bond distance is 1.91 Å and longer Zn-O bond distance is 1.98 Å (see Figure 1(a)) which are in good agreement with the previous theoretical studies.[49,50] The ground state structures of the biomolecules ATP, ADP, and AMP and the methyl phosphate (MP) are shown in Figs. 1(b)-(e) respectively. The numbering of the atoms of the biomolecules and MP along with the bond lengths of the connecting groups of the biomolecules and MP are also shown in the figures. For ATP (Fig.1(b)), the C-N bond lengths in adenine group varies in the range 1.32-1.38 Å, while the bond length of both the N-H bonds is 1.01 Å. C-H bond lengths in ATP vary from 1.08 to 1.1 Å. All the O-H bond lengths are similar with an average value of 0.99 Å. Similarly, the bond lengths of P-O bonds in the phosphate group are in the range 1.49-1.7 Å. The bond lengths for the other biomolecules (ADP (Fig.1(c)) and AMP (Fig.1(d))) are similar as those in ATP. Similarly, in MP (Fig.1(e)), the C-H, P-O and O-H bond lengths are similar as those in the biomolecules.

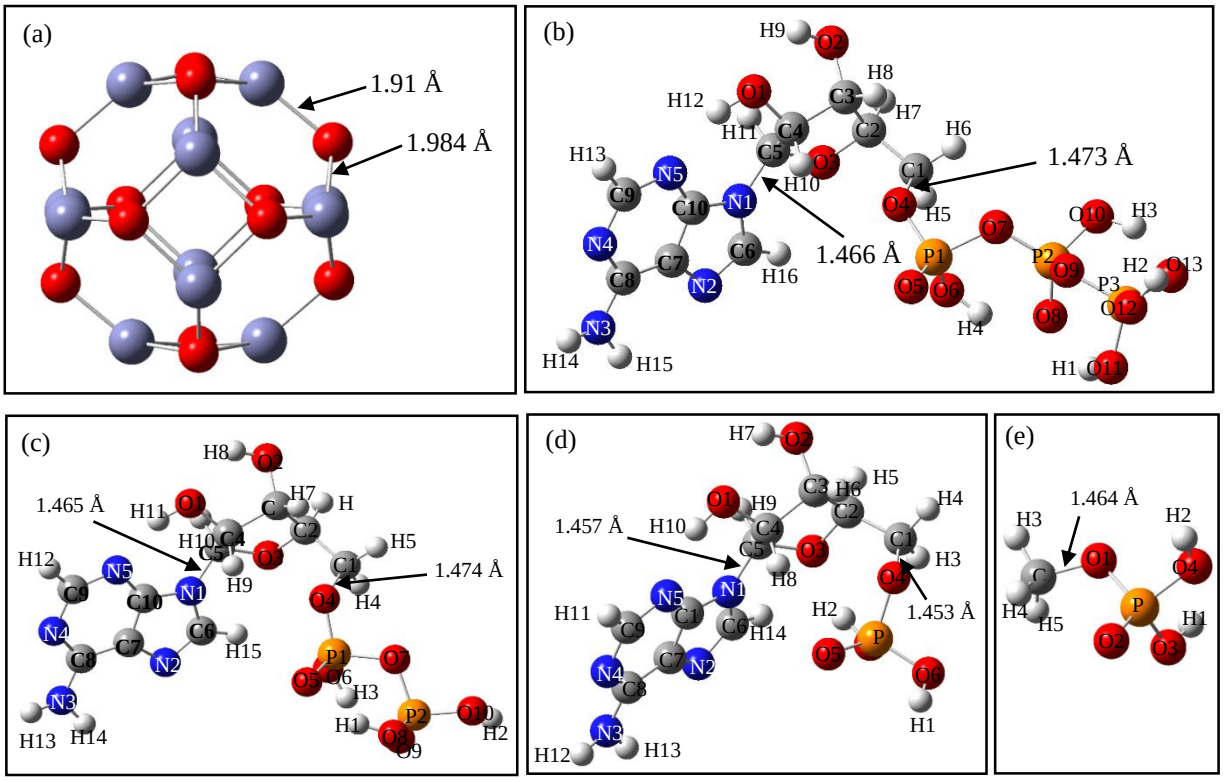

Fig. 1. Optimized geometries of (a) $Zn_{12}O_{12}$ cluster, (b) ATP, (c) ADP, (d) AMP and (e) MP. The bond lengths of each connecting group are given.

Optimized geometries of the nano-bio complexes, along with the bond lengths and the Wiberg bond indices (shown in green) are shown in Fig. 2(a)-(d). ATP forms 5 bonds with ZnONP (Fig. 2(a)), one with N atom of adenine group, Zn1-N2 (2.129 Å) and other four bonds of lengths around 2.0-2.1 Å with oxygens of three phosphate groups, as shown by Zn2-O5, Zn3-O6, Zn4-O8, and Zn4-O11. Non-zero Wiberg bond indices (~ 0.2), as shown in Fig. 2(a), of all these bonds suggest existences of above said bonding interactions in ATP-ZnONP complex. In ADP-ZnONP complex (Fig. 2(b)), ZnONP forms bonds with one N atom of the adenine group (see Zn1-N2 bond (2.116 Å) in Fig. 2(b)) and three bonds of lengths around 2.0 Å with O atoms of two phosphate groups, as shown by Zn2-O5, Zn3-O6 and Zn4-O9 bonds. ZnONP forms one bond Zn1-N2 of length 2.105 Å with the N atom of adenine group of AMP and one bond Zn3-O6 of bond length around 2.0 Å with O atom of the phosphate group, as shown in Fig. 2(c). Zn-N bond distance is similar (~ 2.1 Å) in all three complexes, indicating

that the bonding interaction between ZnO cluster and adenine group does not influence much with increasing number of phosphate groups in bio-ligand from 1 (in AMP) to 3 (in ATP). Note that the ribose group does not participate in bonding with ZnONP in any of these three complexes. ZnONP forms two bonds of length around 2.1 Å with the MP ligand via two oxygen atoms O1 and O2 of the phosphate group, as shown by Zn2-O3 and Zn4-O2 in Fig. 2(d). These Zn-O(phosphate) bond distances are similar to those Zn-O(phosphate) bond distances in ATP-ZnONP, ADP-ZnONP, and AMP-ZnONP complexes indicating that presences of adenine and ribose groups in nano-bio complexes do not influence much the bonding interactions between ZnO cluster and phosphate group. Wiberg bond indices for all complexes, as shown in Fig. 2, also support the bonding interactions between bio ligands and nano particle described above.

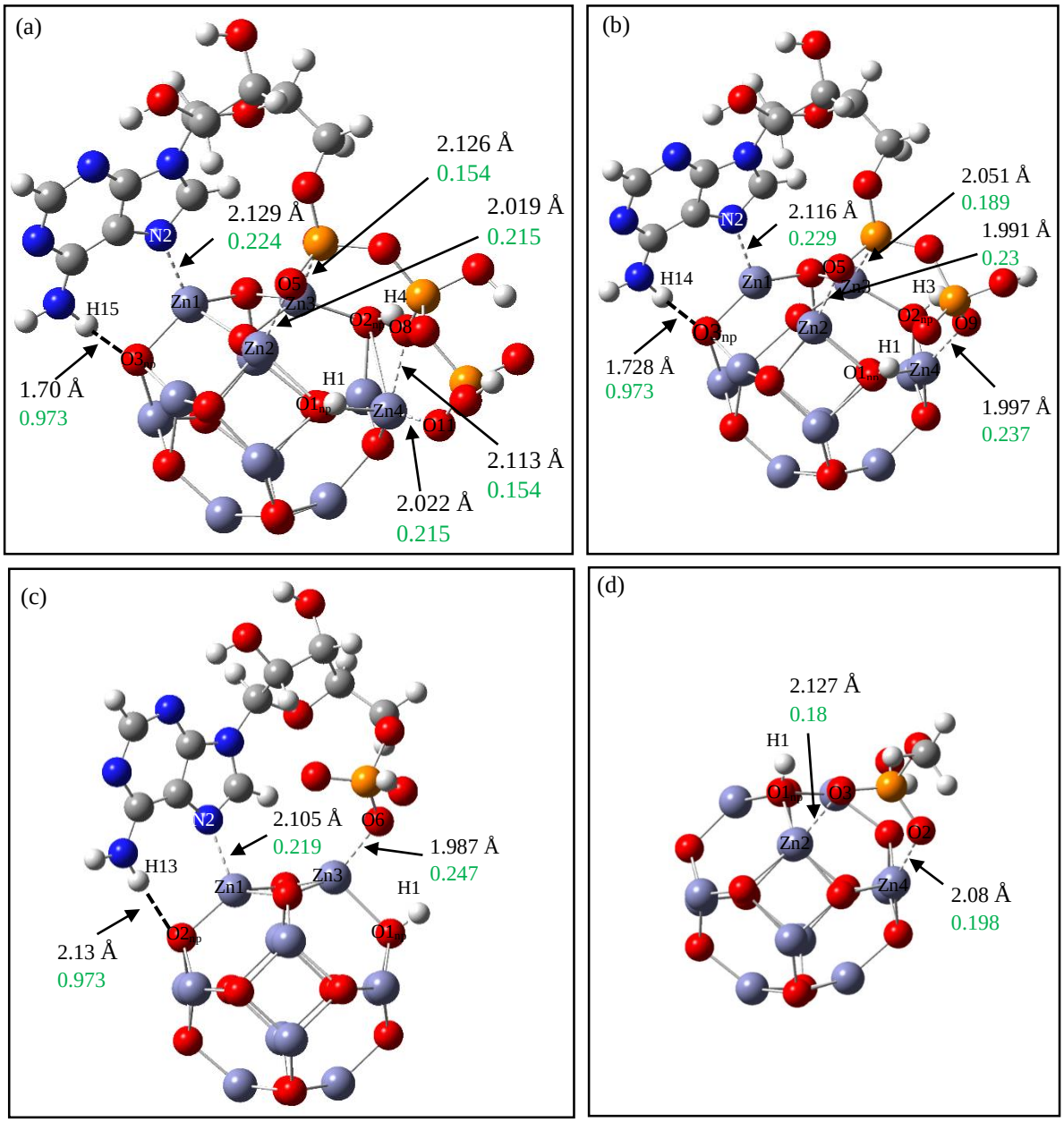

Fig. 2. Optimized geometries of the complexes: (a) ATP-ZnONP, (b) ADP-ZnONP, (c) AMP-ZnONP and (d) MP-ZnONP. The bond lengths with the Wiberg bond indices (shown in green) of the corresponding bonds are shown in the figures.

In addition to bonding interactions discussed above, another type of bond is observed in the complexes (see Figs. 2 (a)-(d). Two H atoms connected with O atoms of two phosphate groups in ATP and ADP ligands form O-H bonds of lengths 0.9 Å with O atoms of ZnO cluster, while two O(phosphate)-H bonds in bio ligands are elongated in the range from 1.7 Å to 2.9 Å in all

three complexes. For instance, the hydrogen atoms H1 and H4 of the phosphate groups at ATP form two new bonds with oxygen atoms $O1_{np}$ and $O2_{np}$ at ZnONP in ATP-ZnONP complex, as shown in Fig. 2(a). At the same time O11(phosphate)-H1 and O6(phosphate)-H4 bonds in ATP are elongated to 2.8 Å and 2.9 Å, respectively. Wiberg bond indices of 0.7 for these new O-H bonds indicate presence of strong O(ZnO)-H bonding interactions in all four complexes. Therefore, two non-classical O-H---O bonding interactions are formed in ATP-ZnONP (Fig. 2(a)) and ADP-ZnONP (Fig. 2(b)) complexes. Such non-classical bonding interaction of H atom is also reported in interaction of transition metal and organosilicon ligand. [51] In AMP-ZnONP complex, only one such non-classical O-H---O bond is formed (see Fig. 2(c)) indicating that number of O-H---O bonds increases with number of phosphate groups in ligands. Presence of O-H---O bond in MP-ZnONP complex as well, shown in Fig. 2(d), indicates that phosphate group interacts with ZnONP through formations of such non-classical bonds. Presence of adenine and ribose groups in ATP, ADP and AMP does not influence much these non-classical O-H---O bond formations. All these geometrical features suggest presence of Zn-N(adenine) and multiple Zn-O(phosphate) bonding interactions along with presence of O-H---O non-classical bonds in all three nano-bio complexes.

Table 1. Binding energies and HOMO-LUMO gaps of the complexes.

Systems	Binding energy, E_B (kcal/mol)	HOMO (eV)	LUMO (eV)	HOMO-LUMO gap, Δ (eV)
Pristine ZnONP	-----	-7.32	-3.37	3.95
ATP-ZnONP	-119.20	-6.75	-3.08	3.67
ADP-ZnONP	-105.90	-6.73	-3.23	3.50
MP-ZnONP	-37.13	-6.80	-3.33	3.47
AMP-ZnONP	-51.50	-6.51	-3.51	3.00

The binding energies (E_B) (see Computational details) of nano-bio complexes are shown in Table 1. The order of E_B is ATP-ZnONP (119.2 kcal/mol) > ADP-ZnONP (105.9

kcal/mol) > AMP-ZnONP (51.5 kcal/mol) > MP-ZnONP (37.13 kcal/mol). Thus, the binding energy increases with increase of number of phosphate group in those molecules. E_B of AMP-ZnONP (51.5 kcal/mol) is higher than that of MP-ZnONP (37.13 kcal/mol) which suggests that presences of adenine groups are also enhances the adsorption of bio molecules onto the metal oxide NP surface.

4.3.2 Important molecular orbitals (MO) of the complexes

Let us now consider different important MO. The Frontier molecular orbitals (FMO) consist of the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO) which provide information on stability and reactivity of a molecule.[52,53] HOMO acts as electron donor and LUMO as electron acceptor. HOMO of ZnONP clusters (Fig.3(a)) mainly consists of O-2p and Zn-3d orbitals, while LUMO (Fig. 3(b)) consists of O-2s and Zn-4s orbitals. HOMO of ATP (Fig.3(c)) is distributed over purine ring of the adenine moiety, while LUMO (Fig3(d)), over phosphate groups. We consider the HOMOs and LUMOs of ADP and AMP and MP as well. Here, both HOMOs and LUMOs of ADP (Fig.3(e) and (f) respectively) and AMP (Fig.3 (g) and (h) respectively) are distributed mainly over purine ring of the adenine moiety. In MP, HOMO (Fig.3 (i)) and LUMO (Fig.3(j)) are distributed all over the molecule.

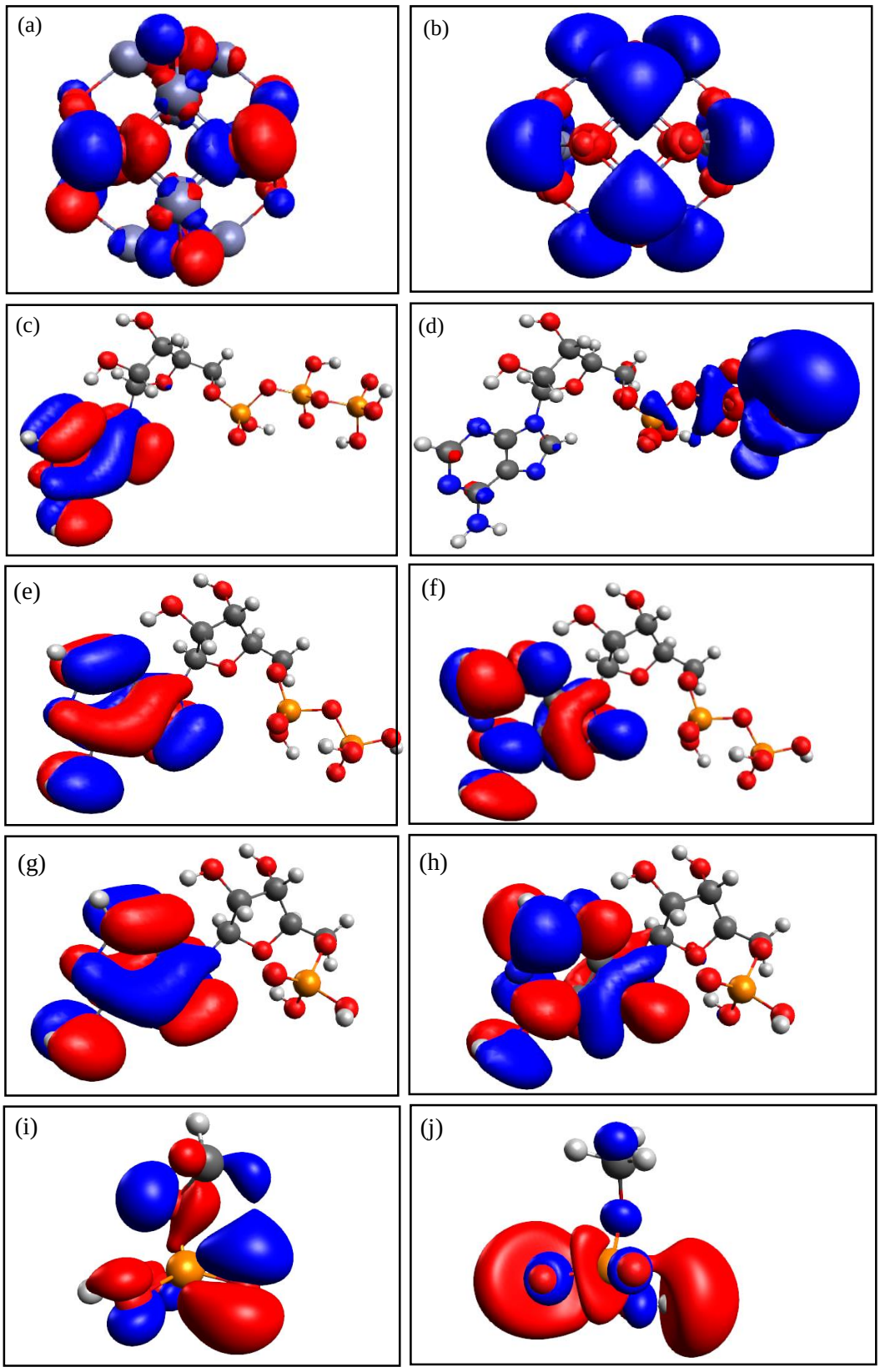

Fig. 3. (a) HOMO and (b) LUMO of $Zn_{12}O_{12}$; (c) HOMO and (d) LUMO of ATP; (e) HOMO and (f) LUMO of ADP; (g) HOMO and (h) LUMO of AMP; (i) HOMO and (j) LUMO of MP.

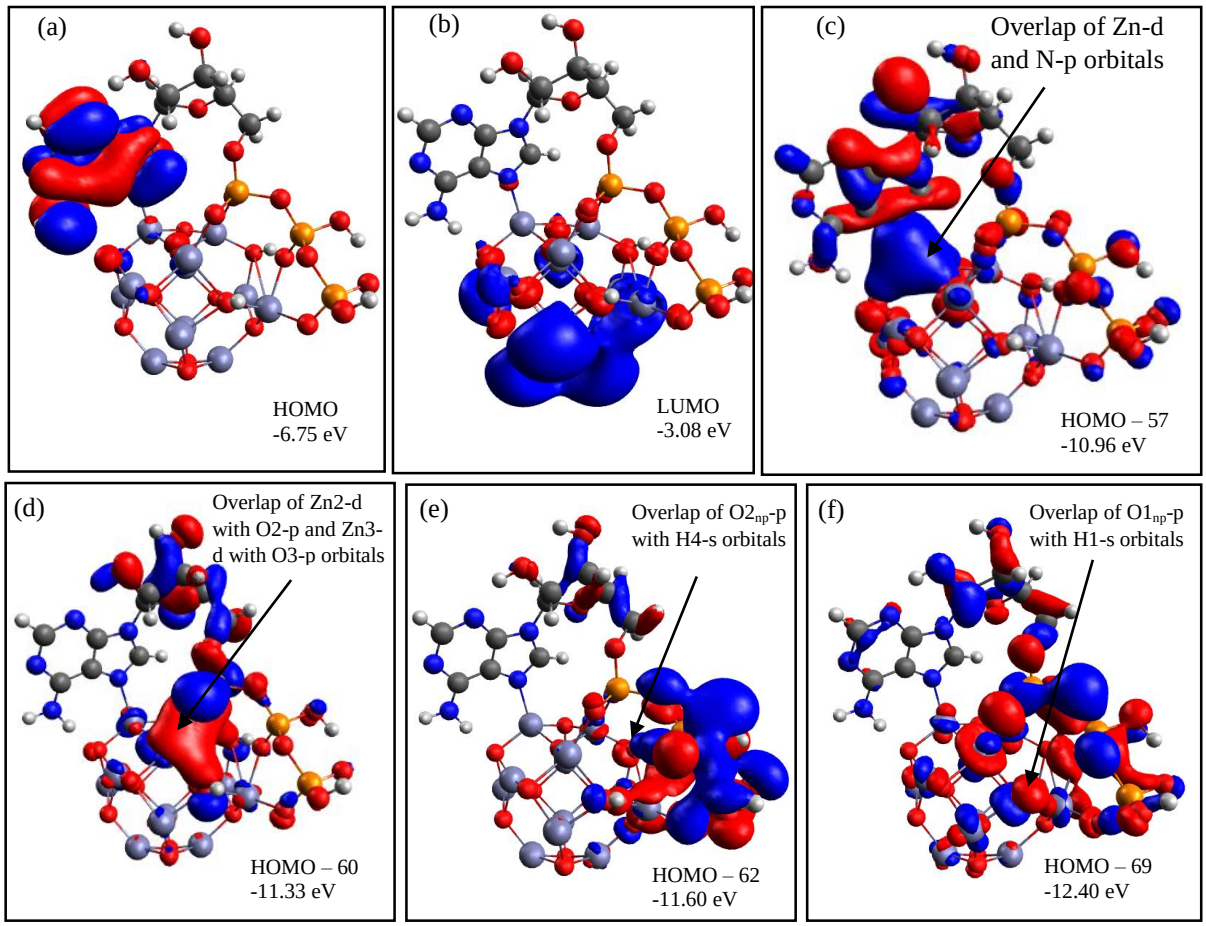

Fig. 4. (a) HOMO, (b) LUMO, (c) HOMO-57, (d) HOMO-60, (e) HOMO-62, (f) HOMO-69 respectively of ZnONP-ATP. The corresponding energy of each state is shown in the figure.

Important Kohn-Sham MOs in ATP-ZnONP are shown in Fig. 4(a)-(f). HOMO (Fig. 4(a)) is mainly distributed over purine ring of adenine, while LUMO (Fig. 4(b)) is located over ZnO cluster. We also consider some MOs near HOMO. The orbital of energy -10.96 eV (HOMO-57, Fig. 4(c)) shows overlap of electron densities between the adenine moiety of ATP and ZnO cluster, while that with energy -11.33 eV (HOMO-60, Fig. 4(d)) shows bonding overlaps between the phosphate groups and ZnO cluster. Two more orbitals with energy -11.60 eV (HOMO-62, Fig. 4(e)) and -12.40 eV (HOMO-69, Fig. 4(f)) represent two O-H bonding overlaps in ATP-ZnONP. Important MOs for ADP-ZnONP are shown in Figs. 5(a)-(f). HOMO (Fig. 5(a)) is distributed over adenine group, while LUMO (Fig. 5(b)) is distributed over the

cluster. HOMO-55 with energy -10.95 eV (see Fig. 5(c)) illustrates the bonding overlap of electron densities between Zn-d orbital and N-p orbital, while HOMO-57 with energy -11.19 eV (see Fig. 5(d)) represents the bonding overlap of phosphate groups with the cluster. The O-H bonding overlaps are represented by HOMO-68 and HOMO-69 with energy -12.59 eV and -12.63 eV respectively, shown in Figs. 5(e) and (f) respectively. Similarly, the MOs of HOMO and LUMO in AMP-ZnONP are same as those in ATP-ZnONP or ADP-ZnONP, shown in Fig. 6(a) and (b). The electron density overlaps between adenine group and the cluster is represented by the orbital of energy -10.51 eV (HOMO-51, Fig. 6(c)), while the orbital of energy -10.62 eV (HOMO-52, Fig. 6(d)) illustrates the bonding overlap between phosphate group and the $Zn_{12}O_{12}$ cluster. The O-H bonding overlap is represented by the orbital of energy -12.77 eV (HOMO-70, Fig. 6(e). The HOMO and LUMO for MP-ZnONP are shown in Fig. 7 (a) and (b) respectively. HOMO is distributed mainly over O atoms of the cluster, while LUMO is located mainly over Zn atoms of the cluster. The bonding overlaps between phosphate group and ZnONP are observed for orbitals with energy -10.4 eV (HOMO-37, Fig. 7 (c)) and -11.39 eV (HOMO-41, Fig. 7(d)).

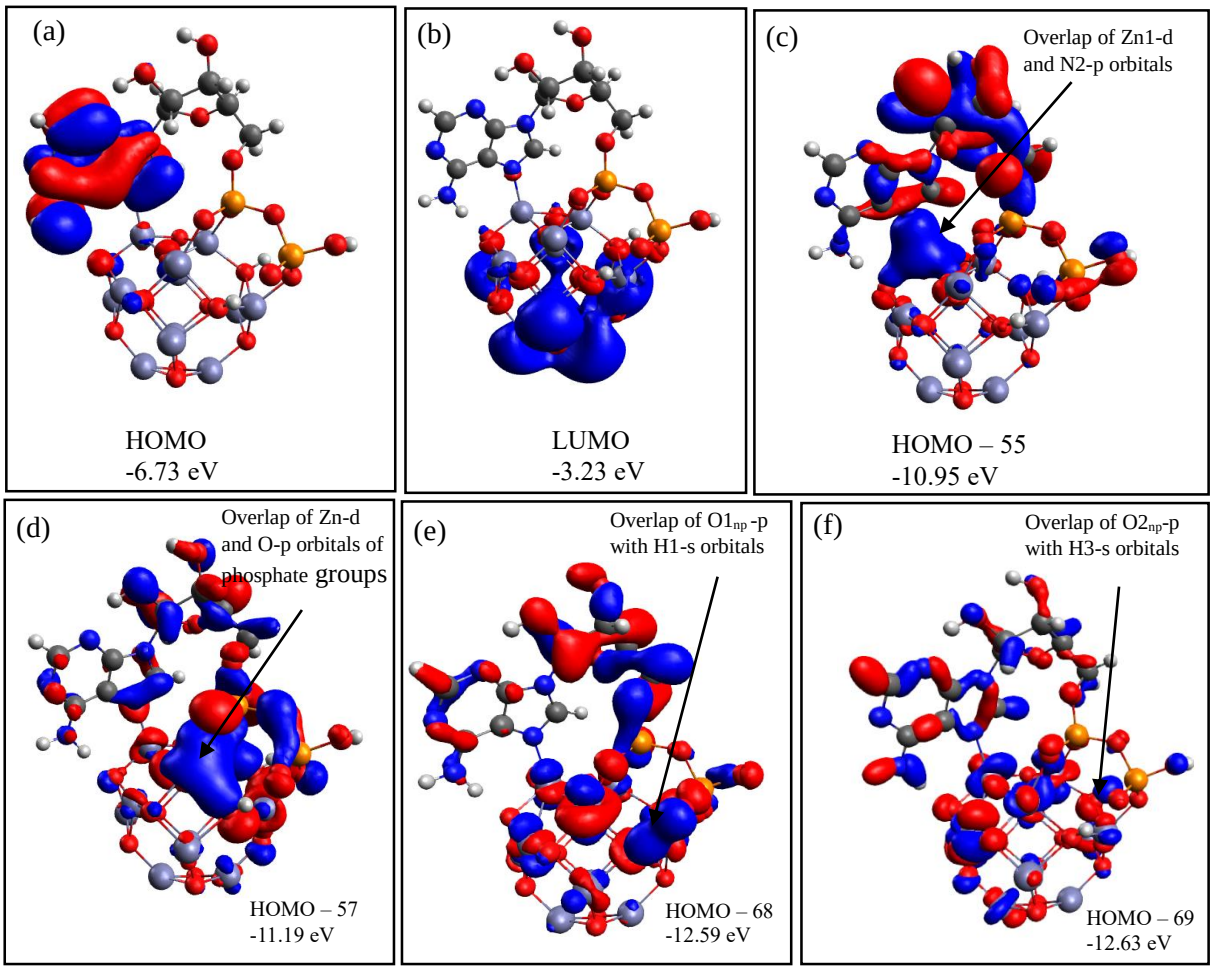

Fig. 5. (a) HOMO, (b) LUMO, (c) HOMO-55, (d) HOMO-57, (e) HOMO-68 and (f) HOMO-69 respectively of ADP-ZnONP.

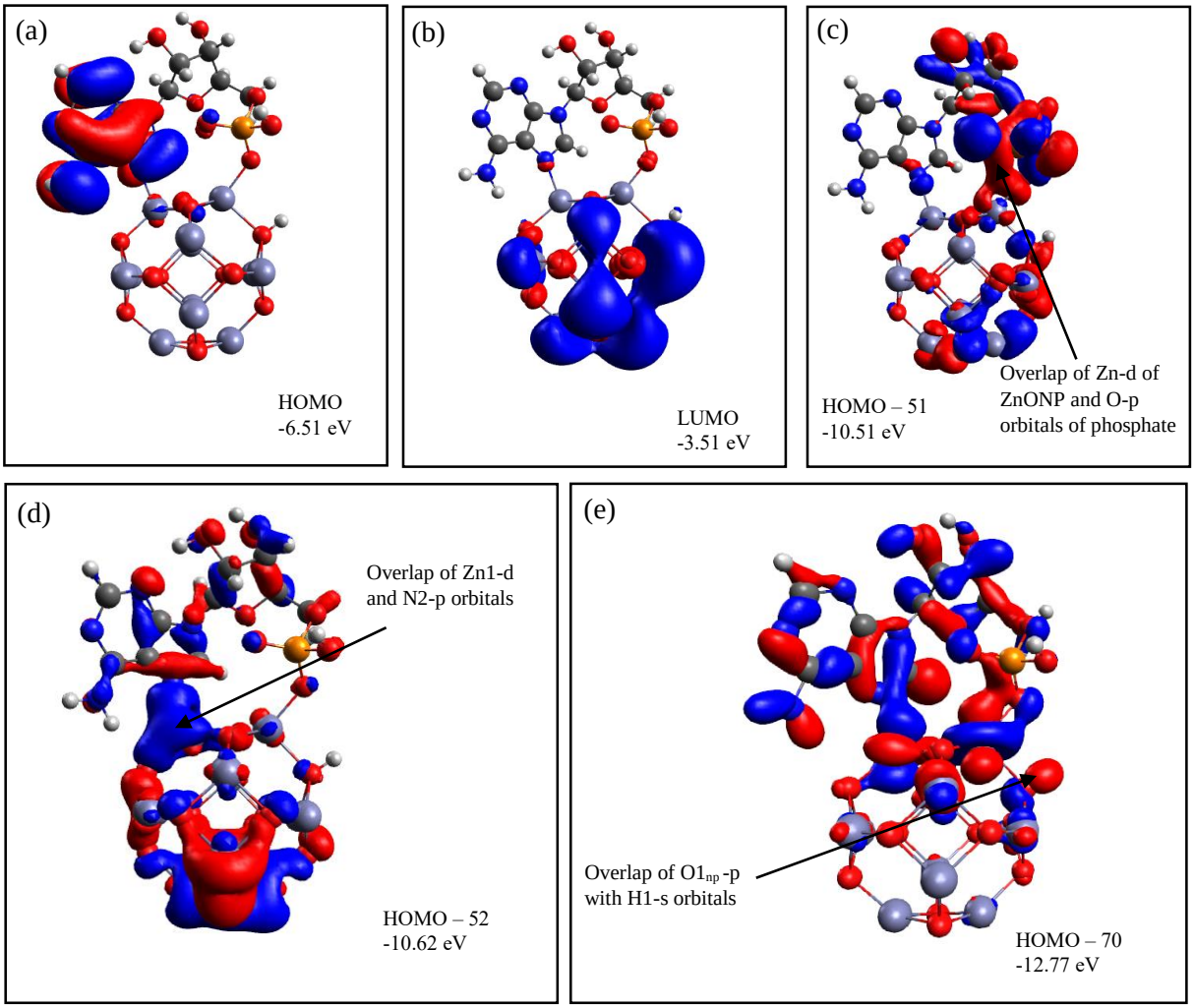

Fig. 6. (a) HOMO, (b) LUMO, (c) HOMO-51, (d) HOMO-52 and (e) HOMO-70 of AMP-ZnONP.

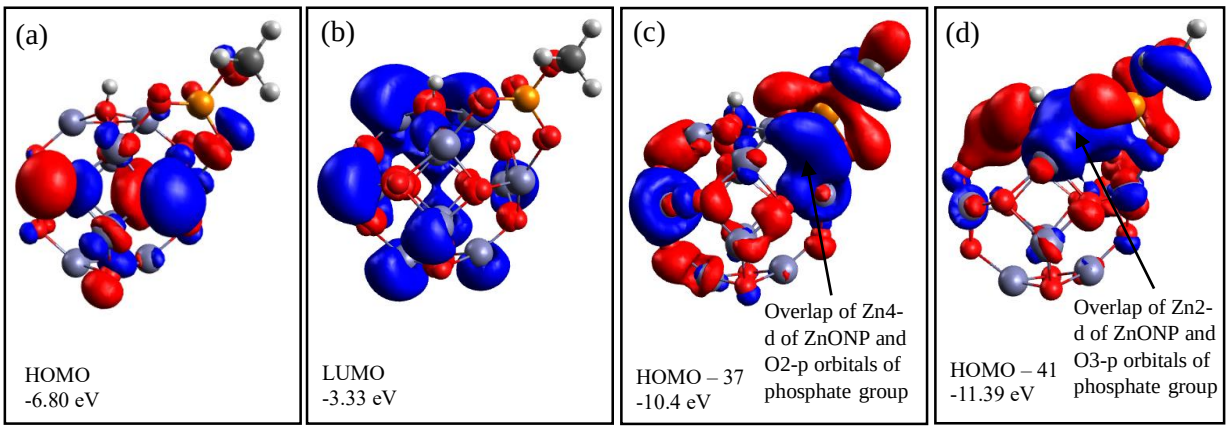

Fig. 7. (a) HOMO, (b) LUMO, (c) HOMO-37 and (d) HOMO-41 of MP-ZnONP.

4.3.3 HOMO-LUMO energy gap and charge transfer (CT)

The HOMO-LUMO energy and the Gap (Δ) between the HOMO-LUMO energy levels for isolated molecules as well as the nano-bio complexes are shown in Table 1. In all nano-bio complexes, Δ value decreases compared to that in pristine ZnONP. The order of Δ is ATP-ZnONP > ADP-ZnONP > MP-ZnONP > AMP-ZnONP. This indicates that the stability of nano-bio complexes increases with increasing number of phosphate group in the ligand. The difference of Δ between MP-ZnONP and AMP-ZnONP again suggests the involvement of the adenine group in the stability.

We illustrate the role of phosphate and adenine groups in the CT. Here, we use natural population analysis (NPA) method to calculate the atomic charges. We calculate atomic charges of each atom of a molecule in free state as well as when it is in the complex. Finally, CT is calculated by subtracting the total atomic charges of each atom in free state from the complex state. Natural populations for isolated molecules and nano-bio complexes are summarized in Table 2. Electron populations of ZnONP decrease in nano-bio complexes than those in isolated ZnONP by 0.42e, 0.37e, 0.18e, and 0.22e in ATP-ZnONP, ADP-ZnONP, AMP-ZnONP, and MP-ZnONP complexes, respectively. In MP-ZnONP complex, population of phosphate group increases by 0.22e than that in isolated MP, while population of ZnONP decreases by the same amount 0.22e from that in isolated ZnONP. These results suggest that electron transfers occur from the ZnONP to phosphate ligand that results in Zn-O coordinate bond formation in MP-ZnONP.

Table 2. Charge transfer (in terms of magnitude of electronic charge *e*) between ZnONP and the biomolecules using natural population analysis (NPA).

Individual contribution	ATP-ZnONP	ADP-ZnONP	MP-ZnONP	AMP-ZnONP
$Zn_{12}O_{12,}\ \Delta q_{NP}$	-0.42	-0.37	-0.22	- 0.18
Phosphate Δq_P	0.49	0.47	0.22	0.27
Adenine Δq_A	-0.07	-0.08	-	−0.07
Ribose Δq_R	-0.00	-0.02	-	-0.02

Let us now consider the presence of both adenine and phosphate groups. Electron populations of adenine groups in nano-bio complexes ATP-ZnONP, ADP-ZnONP and AMP-ZnONP decrease by ~ 0.07e than those in isolated ATP, ADP and AMP molecules (see Table 2), indicating that electron donation occurs from the adenine moiety to ZnONP in all nano-bio complexes. This is supported by presence of Zn-N bonding interaction and electron density overlap between adenine group and ZnONP.

Note that both of MP-ZnONP and AMP-ZnONP complexes have only one phosphate group. The population difference (0.22e) of ZnO cluster between MP-ZnONP complex and isolated ZnO cluster is exactly equal to that (0.22e) of phosphate group between MP-ZnONP and isolated MP but the population difference (0.18e) of ZnO cluster between AMP-ZnONP complex and in isolated ZnO cluster is larger by 0.1e than that (0.28e) of phosphate group in MP-ZnONP and in MP. This is because AMP-ZnONP complex consists electron donations from ZnONP to phosphate group of AMP ligand as well as electron donations from adenine moiety of AMP ligand to ZnONP. Finally, both CTs, ZnONP → phosphate and adenine → ZnONP play major role to stabilize AMP-ZnONP complex through Zn-N and Zn-O(phosphate) coordinate bonds.

The CT results are supported by the corresponding optimized geometry and the MOs. Presence of multiple phosphate groups in bio ligands as in ADP and ATP will therefore increase ZnONP → phosphate CTs which is well reflected in decrease of electron population

of ZnO cluster in ADP-ZnONP and ATP-ZnONP complexes than that in AMP-ZnONP complex where single phosphate group is present (see Table 2). Thus, adenine moiety forms a coordinate covalent bond (Zn-N) with ZnONP through adenine → ZnONP CT and phosphate groups from coordinate covalent bonds (Zn-O) through ZnONP → phosphate CT in all nano-bio complexes. These two types of CTs from ZnO nanocluster to bio-ligand and from bio ligand to ZnO nano-cluster stabilize the nano bio complexes. Multiple Zn-O bonds form in presence of multiple phosphate groups in bio-ligands that increases the stability of nano-bio complexes ADP-ZnONP and ATP-ZnONP compared to that of AMP-ZnONP as ZnONP → phosphate CT increases. Electron population of ribose group does not change much in nano-bio complexes from that in isolated bio molecules as ATP, ADP and AMP which is consistent with absences of bonds as well as absences of electron density overlap between ribose moiety and ZnONP in optimized geometry and MOs discussed above.

Table 3. The changes of the Raman peaks of ZnONP due to the adsorption of ATP, ADP, AMP and MP respectively. Here, ST – stretching, BD – bending. The positive difference means blue shifting, while negative difference, red shifting.

Molecular Vibrations	Wavelength in free ZnONP (cm^{-1})	Shift in wavelength of ZnONP in			
		ATP-ZnONP (cm^{-1})	ADP-ZnONP (cm^{-1})	AMP-ZnONP (cm^{-1})	MP-ZnONP (cm^{-1})
Second order vibrations	73.05	2.72	2.96	5.26	1.65
E2 (low)	97.40	-0.29	-8.11	-2.64	0.88
Second order vibrations	191.40	-2.54	6.11	-11.21	-5.13
Second order vibrations	217.64	-4.71	-3.75	-3.84	-1.44
E1 (TO)	407.47	-1.06	-3.56	0.55	-9.14
Second order vibrations	660.73	4.44	9.36	-1.25	2.09

4.3.4 Changes in Raman frequencies

The wurtzite structure of ZnO belongs to the space group C_{6v}^4, having two formula units per primitive cell, where all atoms are occupying C_{3v} sites. A1+E1+2E2 are the Raman active modes in ZnO, predicted by the group theory [54,55]. The modes A1 and E1 are split into longitudinal optical (LO) and transverse optical (TO) components. Similarly, E2 has two components: E2 (high) vibrational mode associated with oxygen atoms and E2 (low) associated with Zn sublattice [55]. In a wurtzite ZnO, the following fundamental modes exist: E2 (low) at 101 cm^{-1}, A1 (TO) at 380 cm^{-1}, E1 (TO) at 407 cm^{-1}, E2 (high) at 437 cm^{-1}, A1 (LO) at 574 cm^{-1} and E1 (LO) at 583 cm^{-1}, shown in Table 3 [54,55]. Here, we observe two fundamental peaks of bare ZnONP at 97.40 cm^{-1} which is due to E2 (low) mode and at 407.47 cm^{-1} due to E1 (TO) mode. The other peaks of the ZnONP, arises due to the second order vibrations.

We calculate the changes in the Raman spectra of ZnONP in complexes compared to the free ZnONP highlighted in Table 3. The peak at 97.4 cm^{-1} is blue shifted by 0.88 cm^{-1} in MP-ZnONP and red shifted by 2.64 cm^{-1}, 8.11 cm^{-1}, and 0.29 cm^{-1} in AMP-ZnONP, ADP-ZnONP, ATP-ZnONP complexes, respectively. The peak at 407.47 cm^{-1} is red shifted by 9.14 cm^{-1}, 3.56 cm^{-1}, and 1.06 cm^{-1} in MP-ZnONP, ADP-ZnONP, and ATP-ZnONP, respectively and blue shifted by 0.55 cm^{-1} in AMP-ZnONP. All these results suggest presence of strong interactions between bio ligands and ZnONP.

Table 4. Vibrational frequencies of isolated ATP and the shifts of frequencies due to the interactions of ATP with ZnONP. A – adenine group, R – ribose group and P – phosphate group(s). The positive difference means blue shifting, while negative difference means red shifting. The stretching and bending of the bonds are labelled as ST and BD respectively. The missing frequencies are marked in dash (-).

ATP in free state (cm^{-1})	ATP in ATP- ZnONP complex (cm^{-1})	Difference (cm^{-1})	Molecular Vibrations
687.48	701.53	14.05	ST and BD of A, R, P
746.63	752.01	5.38	ST and BD of both A and R
820.03	816.21	-3.82	ST and BD of both R and P
866.08	864.98	-1.1	ST and BD of both A and R
873.8	877.02	3.22	ST and BD of both R and P
882.88	885.89	3.01	C6H16 BD of A
926.24	950.23	23.99	ST and BD of A, R and P
1178.69	1174.86	-3.83	ST and BD of P
1282.15	1289.74	7.59	ST and BD of both A and R
1325.27	1336.04	10.77	Same as above
1370.4	1382.28	11.88	Same as above
1451.46	1448.66	-2.8	Same as above
1501.87	1509.16	7.29	Same as above
1528.73	1548.3	19.57	ST and BD of A and BD of CH groups of R
1612.14	1608.68	-3.46	ST and BD of A
1686.92	1643.99	-42.93	ST and BD of NH2 of A
-	2947.24		N3H14 ST, N3H15 BD of A
3041.36	3062.76	21.4	C3H8, C4H10, C5H11 ST of R
3053.74	3052.25	-1.49	Same as above
3073.68	3080.47	6.79	C1H5, C1H6, C2H7 ST, C5H11, C4H10 BD of R
3088.74	3092.94	4.2	Same as above
3154.91	3149.82	-5.09	C1H5, C1H6 ST of R
-	3193.12		O10H3 ST of P
3206.48	3205.09	-1.39	C9H13 ST of A
3317.03	3333.34	16.31	C6H16 ST, N1C6 and N2C6 BD of A
3407.68	3455.66	47.98	O1H12 ST of R
3427.72	-		OH ST of P
3483.40	-		OH ST of P
3535.70	-		OH ST of P
3590.55	3624.29	33.74	N3H14, N3H15 ST of A
3645.62	3655.17	9.55	O2H9 ST of R
3714.05	-		N3H14, N3H15 ST of A
-	3723.16		OH groups ST of P1
3744.85	3741.34	-3.51	O11H2 ST of P
-	3745.13		OH groups ST of P3

We consider now the Raman frequencies of isolated ATP, ADP, AMP and MP and their shifts in frequencies due to interactions with ZnONP. In isolated ATP, the Raman peaks arising from stretching (ST) and bending (BD) of phosphate group are at 687.48, 746.63, 820.03, 926.24 and 1178.69 cm^{-1}, though there are small contributions of ribose and adenine vibrations in some of the peaks, shown in Table 4. The vibrational frequencies ranging from 1282.15 to 1501.87 cm^{-1} come from the adenine and ribose ST and BD modes, while 1528.73 to 1686.92 cm^{-1} represent ST and BD modes adenine. The peak at 1686.92 cm^{-1} indicates bending of NH2 bonds in adenine. The peak at 3317.03 comes due to the ST of C6H16 bond of adenine. The frequency at 3483.40 is due to the ST of O-H group of phosphates. The vibration at 3714.06 cm^{-1} arises due to the ST of NH2 group of adenine.

Table 5. Vibrational frequencies of ADP and the vibrational shifts due to the interactions of ADP with ZnONP.

ADP in free state (cm^{-1})	ADP in ADP-ZnONP complex (cm^{-1})	Difference (cm^{-1})	Molecular Vibrations
696.29	684.6	-11.69	ST and BD of A, R and P
722.3	732.65	10.35	Same as above
754.36	752.59	-1.77	Same as above
817.72	817.26	-0.46	Same as above
864.14	865.68	1.54	Same as above
878.26	888.68	10.42	Same as above
1085.7	1091.31	6.05	Same as above
1240.65	1236.65	-4	ST and BD of both A and R
1273.65	1276.03	2.38	Same as above
1281.97	1288.50	6.53	Same as above
1339.35	1334.95	-4.4	Same as above
1370.43	1380.83	10.4	Same as above
1451.45	1446.42	-5.03	Same as above
1501.7	1508.1	6.4	Same as above
1528.05	1546.94	18.89	Same as above
1612.52	1609.73	-2.79	ST and BD of A
1687.02	1644.58	-42.44	Same as above
-	2994.59		BD of N3H13 and ST of N3H14
3041.34	3058.07	16.73	ST of CH groups of R
3072.69	3062.8	-9.8	Same as above
3095.85	3092.1	-3.75	Same as above
3161.19	3139.72	-21.47	Same as above
3206.23	3206.0	-0.23	ST of C9H12 of A
3320.97	3333.16	12.19	ST of C6H15 and BD of N1C6 and N2C6 of A
-	3336.72		ST OH groups of P2
3406.40	3473.28	66.88	ST of O1H11 of R
3439.33	-		ST OH groups of P2
3590.23	3625.23	35	ST of N3H13 and N3H14 of A
3641.30	3659.05	17.75	ST of O2H8 of R
3713.68	-		Same as above
3748.10	3754.42	6.32	ST of O10H2 of P1

Table 6. Vibrational frequencies of AMP and their shifting due to the interactions of AMP with ZnONP.

Free AMP (cm^{-1})	AMP in AMP-ZnONP complex (cm^{-1})	Change (cm^{-1})	Molecular Vibrations
715.82	713.22	-2.6	Ribose, Adenine ST, OH BD of P (also BD for all)
742.91	745.4	2.49	Same as above
758.49	768.62	10.13	R, A ST, OP ST of P (also BD for all)
877.7	816.95	-60.75	Vibrations of A group
856.52	860.01	3.49	Vibrations of A, R and P groups
910.9	894.25	-16.65	R, A ST, OP ST of P (also BD for all)
1012.69	1021.97	9.28	Adenine, Ribose ST BD
1030.94	1030.37	-0.57	Adenine, Ribose ST BD, Phosphate OH BD
1239.12	1282.42	43.3	Ribose and Adenine ST, BD
1327.1	1331.71	4.61	Adenine, Ribose ST, BD
1371.15	1381.54	10.39	Adenine, Ribose ST, BD
1454.87	1443.01	-11.86	Same as above
1510.45	1504.56	-5.89	Same as above
1530.42	1541.65	11.23	Adenine ST, BD, Ribose BD
1612.96	1618.15	5.19	Adenine ST, BD
3057.28	3040.74	-16.54	C5H9 ST, C4H9 BD of Ribose
3069.42	-		CH ST of Ribose
3077.24	3059.43	-17.81	CH ST of Ribose
3142.91	3122.91	-20	C1H3, C1H4 ST of Ribose
3208.15	3206.07	-2.08	C9H11 ST of Adenine
3292.64	3344.75	52.11	C6H14 ST, N1C6 and N2C6 BD of Adenine
-	3400.99		NH2 ST of Adenine
3426.29	3512.07	85.78	O1H10 ST of Ribose
-	3652.36		O2H7 ST of Adenine
3514.0	-		OH ST of Phosphate
3591.11	3671.41	80.3	NH2 ST of Adenine
-	3739.89		OH ST P
3663.19	-		OH ST of P
3714.61	-		NH2 ST of Adenine
3759.13	3757.43	-1.7	O6H11 ST of Phosphate

Table 7. Vibrational frequencies of MP and the shifts of frequencies due to the interactions of MP with ZnONP.

Free phosphate (cm^{-1})	Phosphate in phosphate- ZnONP (cm^{-1})	Difference (cm^{-1})	Molecular Vibrations
763.3	768.66	5.36	ST of O2P and CO1
885.04	850.12	-34.92	ST O1P, O2P, O3P and O4P
916.74	1050.87	134.13	ST of O3P, O4P
1027.46	1090.81	63.35	ST CO2
1052.77	-		BD of CO1, O2P ST, O4H2
1070.35	-		BD of O3H1 and O4H2
1168.77	1185.95	17.18	ST of CH3 group
1194.05	1214.49	20.44	ST of CH3 group
1335.83	-		ST of OP bonds of P and BD of O3H1
-	1368.37		BD of O16H2 of P
1498.12	1498.33	0.21	ST of CH3 group
1536.45	1523.53	-12.92	ST of CH3 group
1546.94	1532.86	-14.08	BD of CH3 group
3065.45	3060.98	-4.47	ST of CH3 group
3154.66	3167.41	12.75	Same as above
-	3141.93		ST of O4H2 of P
3183.15	3167.41	-15.74	ST of CH3 group
-	3741.03		ST of OH bonds of P
3753.9	-		ST of O4H2 of P

The wave numbers corresponding to different Raman modes in ADP (see Table 5): The peaks at 722.3 and 754.36 cm^{-1} arise from the BD and ST of adenine and ribose groups. The vibrations in pure ADP ranging from 1240.65 to 1528.05 cm^{-1} comes from ST and BD of adenine and ribose. The peak at 1612.52 cm^{-1} is due to the ST and BD of adenine group. The peak at 1687.02 cm^{-1} consists of the ST of C7C8, C7C10 and N3C8 bonds and the BD of N3H13 and N3H14 bonds. The vibrational frequencies in AMP, ranging from 1030.94 to 1530.42 cm^{-1} is due to the BD and ST of both adenine and ribose groups, shown in Table 6. The peaks at 1612.96 cm^{-1} and 1687.38 cm^{-1} are due to the vibrations of adenine group. The peaks at 3208.15 and 3292.64 cm^{-1} come from the ST of C9H11 and C6H14 bonds. The peaks

in MP (see Table 7) at 1168.77, 1194.05, 1498.12, 1536.45 and 1546.94 cm^{-1} indicate the BD of CH3 group. The peak at 885.04 cm^{-1} indicates the ST of O1P, O2P, O3P and O4P bonds. The peaks at 3065.45, 3154.66 and 3183.15 cm^{-1} come from the stretching of CH3 group.

The frequencies of the Raman spectra of the ligands in complexes following their differences with respect to the free ligands are also shown in the tables. The vibrations ranging from 500 to 1700 cm^{-1} are due to the ST or BD or both of adenine or ribose or phosphate or both for all the biomolecules. We focus on the higher energy region (2000 – 4000 cm^{-1}), where the vibrations of each bond are clearly distinguishable. In ATP-ZnONP (Table 4), the blue shift of 16.31 cm^{-1} of the peak at 3317.03 cm^{-1} is due to the stretching of C6-H16 bond and the bending of N1-C6 and N2-C6 bonds of the adenine group, indicates the interactions of nitrogen (N2) atom of adenine group with zinc (Zn6) atom of ZnONP. The new peak at 2947.24 cm^{-1} due to the stretching of N3-H14 and bending of N3-H15 bonds, indicates the bonding interaction between the N-H (N3-H15) bond of adenine moiety and O (O4) atom of ZnONP. Few vibrational frequencies (3427.72, 3483.40, 3535.70 cm^{-1}) in isolated ATP appear due to the stretching or bending or both of O-H bonds of the phosphate groups (P1 and P3) which are missing in ATP-ZnONP due to the dissociations of these O-H bonds in the complex, as described in previous sections. The new peaks at 3723.16 and 3745.13 cm^{-1} arise due to the O-H stretching of the bonds $O1_{np}$-H1 and $O2_{np}$-H4, where $O1_{np}$ and $O2_{np}$ atoms belong to the ZnONP and the H atoms, H1 and H4 are dissociated from P3 and P1 groups. Similar shifts of vibrational frequencies were observed in other nano-bio complexes ADP-ZnONP (Table 5) and AMP-ZnONP (Table 6) for stretching and bending of C-H, N-C, N-H, O-H bonds of adenine and phosphate moieties that clearly indicate that all these bio-ligands ATP, ADP, and AMP form bonding interactions with the ZnONP via adenine and phosphate groups. In MP-ZnONP complex (Table 7), a peak was observed at 3741.03 cm^{-1} due to the stretching of O-H

($O1_{np}$-H1) bond, where O ($O1_{np}$) atom belongs to ZnONP and H (H1) atom, dissociated from the phosphate group, indicates the interaction of phosphate group with ZnONP.

4.4 Discussion

The chelation of phosphate and adenine groups of ATP to chelate ZnO nanoparticles has been reported from Raman spectroscopy.[9] Our electronic structure, and Raman spectra calculations support this picture of chelation in ATP-ZnONP complex. In addition, we generalize the chelation process to other energy carrier molecules. We observe that the phosphate groups from coordinate covalent bonds (Zn-O) through ZnONP → phosphate CT and adenine moiety forms a coordinate covalent bond (Zn-N) with ZnONP through adenine → ZnONP CT in the nano-bio complexes. Phosphate groups form additional O-H---O non-classical bonds with ZnO nano-cluster in all three nano-bio complexes. All the bonds have their spectral signature in Raman lines.

Hydrolysis of ATP is a highly exergonic process where chemical energy stored in the high-energy phosphoanhydride bonds is released by their splitting [56–58]. The product of ATP hydrolysis is ADP and inorganic phosphate, and further hydrolysed of ADP produce energy, AMP and another phosphate group. ATP hydrolysis is initiated by nucleophilic attack at the third phosphate group by a water molecule. The hydroxide ion (OH^-) of the nucleophilic water molecule attacks the phosphate group in free ATP for dissociation of the phosphoanhydride bond. In free ATP, while the HOMO is distributed over purine ring of the adenine moiety, the LUMO is distributed over the phosphate groups (see in Fig. 3(c) and (d)). The last phosphate group (P3), has the maximum contribution, and is favourable for the nucleophilic attack. However, in ATP-ZnONP complex, the LUMO is distributed over ZnONP, and there is no contribution of ATP on LUMO. This suggests that the ATP hydrolysis pathway, and consequently the energy transfer is likely to be affected by the nano-bio conjugation.

4.5 Conclusions

We have theoretically investigated the interactions of energy carrier biomolecules, ATP, ADP, and AMP with $Zn_{12}O_{12}$ cluster surface using the DFT-based quantum calculations. We establish chelation of the metal oxide surface involving Zn-O bond through ZnONP → phosphate CT and Zn-N bond via adenine → ZnONP CT and additional O-H---O non-classical bonds with ZnO nano-cluster with phosphate groups in all the nano-bio complexes. DFT-calculated Raman spectra strongly support presence of these bonds in all three nano-bio complexes. Our results may also be important to understand how the metal oxide nanoparticles affect the cellular energy transfer reactions and the reaction pathways.

4.6 Appendix A

Details of the methods

A.1 Introduction

The time independent Schrodinger equation in one dimension is,

$$\frac{-\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = \hat{H}\psi(x) = E\psi(x) \qquad (1)$$

where m is the mass of the particle, $\hbar = (h/2\pi)$ is the reduced Planck's constant, $\psi(x)$ is the wavefunction, $V(x)$ is the potential energy, $\hat{H}$ is the Hamiltonian operator and E is the total energy. $\psi(x)$ describes the behavior of a particle. The energy of the system E can be obtained by solving the Schrodinger equation. The Schrodinger equation, on the other hand, can only be solved exactly for H atom and $H2^+$ molecule. To solve Schrodinger equation beyond H atom and $H2^+$ molecule, density functional theory (DFT) is one most important and powerful tool that is used widely nowadays.

A.2 Many-body problem

The many-body Hamiltonian of a system of N_e electrons and N_n nuclei is of the form,

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn}$$

$$= \frac{-\hbar^2}{2m_e}\sum_i^{N_e}\nabla_i^2 + \frac{-\hbar^2}{2m_n}\sum_i^{N_n}\nabla_i^2 + \frac{e^2}{4\pi\varepsilon_0}\left[-\sum_i^{N_e}\sum_I^{N_n}\frac{Z_I}{|\vec{r}_i-\vec{r}_I|} + \frac{1}{2}\sum_i^{N_e}\sum_{j.i}^{N_e}\frac{1}{|\vec{r}_i-\vec{r}_j|} + \frac{1}{2}\sum_I^{N_n}\sum_{J.I}^{N_n}\frac{1}{|\vec{r}_I-\vec{r}_J|}\right] \quad (2)$$

where i and j run over the electronic degrees of freedom, and I and J run over the nuclei degrees of freedom. m_e, m_n and Z_I are mass of electron, mass of nuclei and atomic number respectively. $\vec{r}_i$ and $\vec{r}_I$ are the positions of i^{th} electorn and I^{th} nucleus respectively. The Hamiltonian consists of five terms: the kinetic energy of electrons, kinetic energy of nuclei, electron-nuclei interactions, electron-electron interactions and nuclei-nuclei interactions. According to the Born-Oppenheimer approximation, the nuclear motion is neglected as the mass of nucleons is a few orders of magnitude higher than that of electron. So, the nuclei can be considered as classical particles that generates an external potential, and the electrons as quantum particles are subjected to that potential. So, equation (2) can be written as,

$$\hat{H} = \frac{-\hbar^2}{2m_e}\sum_i^{N_e}\nabla_i^2 + \frac{e^2}{4\pi\varepsilon_0}\left[-\sum_i^{N_e}\sum_I^{N_n}\frac{Z_I}{|\vec{r}_i-\vec{r}_I|} + \frac{1}{2}\left[\sum_i^{N_e}\sum_{j.i}^{N_e}\frac{1}{|\vec{r}_i-\vec{r}_j|}\right]\right] \quad (3)$$

A.3 Density functional theory (DFT)

DFT is a very powerul and most commonly used tool to calulate the electronic properties of many-body systems. Thomas and Fermi first propose the initial version of DFT in 1927. Later, DFT of the quantum ground state is established by Hohenberg, Kohn and Sham which is superior to both Thomas-Fermi and Hartree-Fock theories, allowing for a wide range of applications in real-world physical systems. In theoretical physics and molecular chemistry, DFT becomes one of the main tools for electronic structure calculations and many other. In

DFT, the properties of a system are not calculated through the many electron wavefunction, but calculated from the electron density (ρ(r)). The total energy is given by,

$$E\big(\rho(r)\big) = T_e\,[\rho(r)] + V_{ne}\,[\rho(r)] + V_{ee}\,[\rho(r)] + E_{xc}[\rho(r)], \qquad (4)$$

where T_e is the kinetic energy of electron, V_{ne} is the nuclei-electron potential energy, V_{ee} is the electron-electron potential energy, and E_{xc} is the exchange-correlation energy. E_{xc} is the only term unknown here. The estimate of the exchange-correlation energy causes the biggest error in DFT. In order to calculate E_{xc} two most useful approximations are used: local-density approximation (LDA) and generalized gradient approximation (GGA).

A.3.1 Exchange-correlation energy approximations

A.3.1.1 Local-density approximation (LDA)

LDA, an approximation to the exchange-correlation energy, depends only on the electronic density at each point in space. Here, the local density can be treated as a uniform electron gas, i.e., the system is locally homogeneous. For a spin-unpolarized system, the XC energy can be written as

$$E_{xc}^{LDA}(\rho(\boldsymbol{r})) = \int \rho(r)\varepsilon_{xc}^{LDA}\,(\rho(\boldsymbol{r}))d\boldsymbol{r}, \qquad (5)$$

where $\varepsilon_{xc}^{LDA}(\rho(\boldsymbol{r}))$ is the exchange-correlation energy per particle. The systems in which the electron density varies slowly, LDA gives very good results. However, LDA predicts that transition metal oxides XO (X = Fe, Mn, Ni) are either semiconductors or metals, even though they are Mott insulators. LDA gives wrong ground state for many simpler systems, and also the description of hydrogen bonding using this functional is very poor.

A.3.1.2 Generalized gradient approximation (GGA)

As the real systems are inhomogeneous, the variation in electron density over space is taken into account by GGA. Here, the exchange-correlation energy can be written as,

$$E_{xc}^{GGA} = E_{xc}[\rho(r), \nabla \rho(r)] \qquad (6)$$

For molecular geometries, binding energies, and ground-state energies, the GGA functional provides very good results. Nowadays, to deal with the gradient of the electron density, there are many different functionals available, such as PBE (Perdew-Burke-Ernzerhof), B3LYP (Becke, 3-parameter, Lee–Yang–Parr), etc.

B3LYP

B3LYP is a Hybrid functional. Hybrid functionals are a class of approximations to the exchange–correlation energy functional that incorporate a part of exact exchange from Hartree–Fock theory, while the rest of the exchange–correlation energy is incorporated from ab initio or empirical method. The form of the B3LYP functional is,

$$E_{xc}^{B3LYP} = (1-a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + cE_c^{LYP} + (1-c)E_c^{LSDA}, \qquad (7)$$

where $a = 0.20$, $b = 0.72$ and $c = 0.81$. E_x^{LSDA} is the standard local exchange functional of the local spin density approximation, and E_x^{HF} is the exact Hartree–Fock exchange. E_x^{B88} and E_c^{LYP} are Becke's gradient correction to the exchange functional and Lee–Yang–Parr correlation functional, respectively. E_c^{LSDA} is the VWN (Vosko, Wilk, and Nusair) local spin density approximation to the correlation functional.

A.4 Basis sets

A basis set is a set of functions used to describe the electronic wavefunction in the Hartree-Fock method or in DFT. Quantum chemistry community uses basis set as a linear

combinations of atomic orbitals, while the solid-state community uses plane wave basis set. Atomic orbitals can be Gaussian-type orbitals, Slater-type orbitals, etc. The wavefunction in plane wave basis set has the form,

$$\psi_k(\boldsymbol{r}) = \sum_G c_{k+G}\, e^{i(\boldsymbol{k}+\boldsymbol{G}).\boldsymbol{r}}, \qquad (8)$$

where $\boldsymbol{k}$ and $\boldsymbol{G}$ are reciprocal lattice vectors. c_{k+G} are plane waves coefficients.

Gaussian-type orbital (GTO) has the form,

$$\Phi(x, y, z) = N x^a y^b z^c e^{-\xi r^2}, \qquad (9)$$

where N is normalization constant, a, b and c controls the angular momentum, L= $a + b + c$. ξ controls the width of the orbital. We have used 6-31G(d), 6-311++G(d,p) and LANL2DZ basis sets in this chapter.

6-31G(d)

This is a double zeta split valence basis. Here, the core orbitals are described by a sum of six GTOs, while the inner part of the valence orbitals is a contraction of three GTOs, and the outer part of the valence orbitals represented by one GTO. d inside the bracket represents that we include d-type polarization function on heavy atoms.

6-311++G(d,p)

This is a triple zeta split valence basis where the core orbitals are described by the sum of six GTOs. The valence orbitals are split into three functions, described by three, one and one GTOs respectively. The first + indicates that one set of diffuse s and p functions are added to the heavy atoms, while the second + represents that diffuse s function is also added to the hydrogen atoms. d and p illustrate that the d type polarization on heavy atoms and p type polarization function on hydrogen atoms respectively are included.

LANL2DZ (Los Alamos National Laboratory 2 double ζ)

LANL2DZ is an effective core potentials (ECPs) basis used for transition metals. ECPs are a useful means of replacing the core electrons of atomic or molecular systems by an effective potential, while the valance electrons are treated explicitly. As a result, the computational effort is significantly reduced.

A.5 Natural bond orbital (NBO) analysis

Natural bond orbital analysis is a useful tool for understanding hybridization and covalency effects in polyatomic wave functions. NBO optimally transforms a given wave function into localized form which corresponds to the one-center ("lone pairs") and two-center ("bonds") elements of the chemist's Lewis structure picture. Natural atomic orbitals (NAOs) and natural hybrid orbitals (NHOs) are used to transform the input atomic orbital basis set into natural bond orbitals in NBO analysis.

A.6 Gaussian 16 code

Gaussian 16 is the most recent version of the Gaussian electronic structure software series, which is used by chemists, physicists, chemical engineers, biochemists, and other professionals. It predicts the energies, molecular structures, vibrational frequencies of compounds. It not only determines the ground state structures, but also predicts the structures of transition states. Both stable species and chemicals that are difficult or impossible to examine experimentally can be studied from Gaussian 16's models. Gaussian 16 uses different methods for electronic structure calculations including Hartree-Fock, Density functional theory (DFT), semi-empirical: AM1, PM6, PM7 and DFTB, etc.

4.7 References

(1) Phogat, N.; Kohl, M.; Uddin, I.; Jahan, A. Interaction of Nanoparticles with Biomolecules, Protein, Enzymes, and Its Applications. In *Precision Medicine*; Elsevier, 2018; pp 253–276.

(2) Auría-Soro, C.; Nesma, T.; Juanes-Velasco, P.; Landeira-Viñuela, A.; Fidalgo-Gomez, H.; Acebes-Fernandez, V.; Gongora, R.; Almendral Parra, M. J.; Manzano-Roman, R.; Fuentes, M. Interactions of Nanoparticles and Biosystems: Microenvironment of Nanoparticles and Biomolecules in Nanomedicine. *Nanomaterials* **2019**, *9* (10), 1365.

(3) Jeevanandam, J.; Barhoum, A.; Chan, Y. S.; Dufresne, A.; Danquah, M. K. Review on Nanoparticles and Nanostructured Materials: History, Sources, Toxicity and Regulations. *Beilstein J. Nanotechnol.* **2018**, *9* (1), 1050–1074.

(4) Sezer, A. D. *Application of Nanotechnology in Drug Delivery*; BoD–Books on Demand, 2014.

(5) Limo, M. J.; Sola-Rabada, A.; Boix, E.; Thota, V.; Westcott, Z. C.; Puddu, V.; Perry, C. C. Interactions between Metal Oxides and Biomolecules: From Fundamental Understanding to Applications. *Chem. Rev.* **2018**, *118* (22), 11118–11193.

(6) Seabra, A. B.; Durán, N. Nanotoxicology of Metal Oxide Nanoparticles. *Metals* **2015**, *5* (2), 934–975.

(7) Alim, K. A.; Fonoberov, V. A.; Shamsa, M.; Balandin, A. A. Micro-Raman Investigation of Optical Phonons in ZnO Nanocrystals. *J. Appl. Phys.* **2005**, *97* (12), 124313.

(8) Rumyantseva, A.; Kostcheev, S.; Adam, P.-M.; Gaponenko, S. V.; Vaschenko, S. V.; Kulakovich, O. S.; Ramanenka, A. A.; Guzatov, D. V.; Korbutyak, D.; Dzhagan, V.

Nonresonant Surface-Enhanced Raman Scattering of ZnO Quantum Dots with Au and Ag Nanoparticles. *ACS Nano* **2013**, *7* (4), 3420–3426.

(9) Bhaumik, A.; Shearin, A. M.; Delong, R.; Wanekaya, A.; Ghosh, K. Probing the Interaction at the Nano–Bio Interface Using Raman Spectroscopy: ZnO Nanoparticles and Adenosine Triphosphate Biomolecules. *J. Phys. Chem. C* **2014**, *118* (32), 18631–18639.

(10) Ramani, M.; Nguyen, T. D. T.; Aryal, S.; Ghosh, K. C.; DeLong, R. K. Elucidating the RNA Nano–Bio Interface: Mechanisms of Anticancer Poly I: C RNA and Zinc Oxide Nanoparticle Interaction. *J. Phys. Chem. C* **2017**, *121* (29), 15702–15710.

(11) Martinez, G. R.; Loureiro, A. P. M.; Marques, S. A.; Miyamoto, S.; Yamaguchi, L. F.; Onuki, J.; Almeida, E. A.; Garcia, C. C.; Barbosa, L. F.; Medeiros, M. H. Oxidative and Alkylating Damage in DNA. *Mutat. Res. Mutat. Res.* **2003**, *544* (2–3), 115–127.

(12) Sharma, V.; Shukla, R. K.; Saxena, N.; Parmar, D.; Das, M.; Dhawan, A. DNA Damaging Potential of Zinc Oxide Nanoparticles in Human Epidermal Cells. *Toxicol. Lett.* **2009**, *185* (3), 211–218.

(13) Maddahi, P. S.; Yeganeh, M.; Baghsiyahi, F. B. ZnO Nanoparticles as a Sensitive Platform for Detection of Nitration in Tyrosine and Tryptophan: A DFT Study. *Mater. Chem. Phys.* **2019**, *237*, 121857.

(14) Chandraboss, V. L.; Karthikeyan, B.; Senthilvelan, S. Experimental and First-Principles Study of Guanine Adsorption on ZnO Clusters. *Phys. Chem. Chem. Phys.* **2014**, *16* (42), 23461–23475.

(15) Ito, M. In Vitro Properties of a Chitosan-Bonded Hydroxyapatite Bone-Filling Paste. *Biomaterials* **1991**, *12* (1), 41–45.

(16) Ørstavik, D.; Hongslo, J. K. Mutagenicity of Endodontic Sealers. *Biomaterials* **1985**, *6* (2), 129–132.

(17) Gong, M.; Yang, Z.; Xu, X.; Jasion, D.; Mou, S.; Zhang, H.; Long, Y.; Ren, S. Superhydrophobicity of Hierarchical ZnO Nanowire Coatings. *J. Mater. Chem. A* **2014**, *2* (17), 6180–6184.

(18) Lu, G.; Maragakis, P.; Kaxiras, E. Carbon Nanotube Interaction with DNA. *Nano Lett.* **2005**, *5* (5), 897–900.

(19) Nel, A.; Xia, T.; Mädler, L.; Li, N. Toxic Potential of Materials at the Nanolevel. *science* **2006**, *311* (5761), 622–627.

(20) Velegol, D. Assembling Colloidal Devices by Controlling Interparticle Forces. *J. Nanophotonics* **2007**, *1* (1), 012502.

(21) Nel, A. E.; Mädler, L.; Velegol, D.; Xia, T.; Hoek, E. M.; Somasundaran, P.; Klaessig, F.; Castranova, V.; Thompson, M. Understanding Biophysicochemical Interactions at the Nano–Bio Interface. *Nat. Mater.* **2009**, *8* (7), 543.

(22) Min, Y.; Akbulut, M.; Kristiansen, K.; Golan, Y.; Israelachvili, J. The Role of Interparticle and External Forces in Nanoparticle Assembly. In *Nanoscience And Technology: A Collection of Reviews from Nature Journals*; World Scientific, 2010; pp 38–49.

(23) Furmaniak, S.; Terzyk, A. P.; Jaroniec, M.; Gauden, P. A. The Chemistry of Bioconjugation in Nanoparticles-Based Drug Delivery System. *Adv. Condens. Matter Phys.* **2015**, *2015* (1), 1–27.

(24) Kumar, V. S. P.; Verma, M.; Deshpande, P. A. On Interaction of Arginine, Cysteine and Guanine with a Nano-TiO2 Cluster. *Comput. Biol. Chem.* **2020**, 107236.

(25) Whatley, F. R.; Allen, M. B.; Trebst, A. V.; Arnon, D. I. Photosynthesis by Isolated Chloroplasts IX. Photosynthetic Phosphorylation and CO2 Assimilation in Different Species. *Plant Physiol.* **1960**, *35* (2), 188.

(26) Morrison, J. F.; Heyde, E. Enzymic Phosphoryl Group Transfer. *Annu. Rev. Biochem.* **1972**, *41* (1), 29–54.

(27) Wood, H. G.; o'brien, w. e. properties of carboxytransphosphorylase; pyruvate, phosphate dikinase; pyrophosphate-phosphofructokinase and pyrophosphate-acetate kinase and their roles in the metabolism of inorganic pyrophosphate. **1977**.

(28) Eichhorn, G. L. *Inorganic Biochemistry*; Elsevier, 1973; Vol. 2.

(29) Aggarwal, A. K. Structure and Function of Restriction Endonucleases. *Curr. Opin. Struct. Biol.* **1995**, *5* (1), 11–19.

(30) Mandal, S. C.; Maganti, L.; Mondal, M.; Chakrabarti, J. Microscopic Insight to Specificity of Metal Ion Cofactor in DNA Cleavage by Restriction Endonuclease EcoRV. *Biopolymers* e23396.

(31) Lanir, A.; Yu, N.-T. A Raman Spectroscopic Study of the Interaction of Divalent Metal Ions with Adenine Moiety of Adenosine 5'-Triphosphate. *J. Biol. Chem.* **1979**, *254* (13), 5882–5887.

(32) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785.

(33) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F. N.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627.

(34) Ray, M.; Nakao, Y.; Sato, H.; Sakaba, H.; Sakaki, S. Silapropargyl/Silaallenyl and Silylene Acetylide Complexes of [Cp (CO) 2W]+. Theoretical Study of Their Interesting Bonding Nature and Formation Reaction. *J. Am. Chem. Soc.* **2006**, *128* (36), 11927–11939.

(35) Ray, M.; Nakao, Y.; Sato, H.; Sakaki, S. Theoretical Study of Tungsten H3-Silaallyl/H3-Vinylsilyl and Vinyl Silylene Complexes: Interesting Bonding Nature and Relative Stability. *Organometallics* **2007**, *26* (18), 4413–4423.

(36) Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*; Wavefunction Irvine, CA, 2003; Vol. 2.

(37) Carrasco, J.; Illas, F.; Bromley, S. T. Ultralow-Density Nanocage-Based Metal-Oxide Polymorphs. *Phys. Rev. Lett.* **2007**, *99* (23), 235502.

(38) Matxain, J. M.; Mercero, J. M.; Fowler, J. E.; Ugalde, J. M. Electronic Excitation Energies of Zn i O i Clusters. *J. Am. Chem. Soc.* **2003**, *125* (31), 9494–9499.

(39) Yong, Y.; Song, B.; He, P. Growth Pattern and Electronic Properties of Cluster-Assembled Material Based on Zn12O12: A Density-Functional Study. *J. Phys. Chem. C* **2011**, *115* (14), 6455–6461.

(40) Reber, A. C.; Khanna, S. N.; Hunjan, J. S.; Beltran, M. R. Rings, Towers, Cages of ZnO. *Eur. Phys. J. D* **2007**, *43* (1–3), 221–224.

(41) Dmytruk, A.; Dmitruk, I.; Shynkarenko, Y.; Belosludov, R.; Kasuya, A. ZnO Nested Shell Magic Clusters as Tetrapod Nuclei. *RSC Adv.* **2017**, *7* (35), 21933–21942.

(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. Gaussian 16. *Revis. A* **2016**, *3*.

(43) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82* (1), 270–283.

(44) Wadt, W. R.; Hay, P. J. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. *J. Chem. Phys.* **1985**, *82* (1), 284–298.

(45) Wedig, U.; Dolg, M.; Stoll, H.; Preuss, H. Energy-Adjusted Pseudopotentials for Transition-Metal Elements. In *Quantum chemistry: the challenge of transition metals and coordination chemistry*; Springer, 1986; pp 79–89.

(46) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint. *Chem. Rev.* **1988**, *88* (6), 899–926.

(47) Glendening, E. D.; Landis, C. R.; Weinhold, F. Natural Bond Orbital Methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 1–42.

(48) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. Doi: 10.1021/Ja980917m. *J Am Chem Soc* **1998**, *120*, 12051.

(49) Zhao, M.; Xia, Y.; Tan, Z.; Liu, X.; Mei, L. Design and Energetic Characterization of ZnO Clusters from First-Principles Calculations. *Phys. Lett. A* **2007**, *372* (1), 39–43.

(50) Wang, B.; Nagase, S.; Zhao, J.; Wang, G. Structural Growth Sequences and Electronic Properties of Zinc Oxide Clusters (ZnO) n (N= 2-18). *J. Phys. Chem. C* **2007**, *111* (13), 4956–4963.

(51) Whited, M. T.; Taylor, B. L. Metal/Organosilicon Complexes: Structure, Reactivity, and Considerations for Catalysis. *Comments Inorg. Chem.* **2020**, *40* (5), 217–276.

(52) Fukui, K.; Yonezawa, T.; Shingu, H. A Molecular Orbital Theory of Reactivity in Aromatic Hydrocarbons. *J. Chem. Phys.* **1952**, *20* (4), 722–725.

(53) Fleming, I. *Frontier Orbitals and Organic Chemical Reactions*; Wiley, 1977.

(54) Damen, T. C.; Porto, S. P. S.; Tell, B. Raman Effect in Zinc Oxide. *Phys. Rev.* **1966**, *142* (2), 570.

(55) Arguello, C. A.; Rousseau, D. L.; Porto, S. da S. First-Order Raman Effect in Wurtzite-Type Crystals. *Phys. Rev.* **1969**, *181* (3), 1351.

(56) Lodish, H.; Zipursky, S. L. Molecular Cell Biology. *Biochem Mol Biol Educ* **2001**, *29*, 126–133.

(57) Rosing, J.; Slater, E. C. The Value of ΔG for the Hydrolysis of ATP. *Biochim. Biophys. Acta BBA-Bioenerg.* **1972**, *267* (2), 275–290.

(58) Cox, M. M.; Nelson, D. L. *Lehninger Principles of Biochemistry*; Wh Freeman New York, 2008; Vol. 5.

CHAPTER 5

Surface Specific Adsorption of Glucose to ZnO

5.1 Introduction

In the previous chapter we discuss the chelation of energy carrier biomolecules with ZnO nanoparticles. ZnO acts as a promising candidate [10,11] for bio- sensors. Although glucose is the most important source of energy in all organisms, an abnormal level of this substance in human body can lead to heart failure, stroke, diabetes, obesity and so on. Monitoring glucose levels in human blood is necessary to avoid such diseases. Even in the manufacturing sectors, such as food, beverage and fermentation, the control of glucose level is important to maintain the quality and safety of ingredients. Hence, a sensitive, fast and reliable biosensor is necessary for detection of glucose levels.

Glucose biosensors have been developed based on the immobilization of glucose oxidase (GO) enzyme on different matrices.[13,14] The enzyme-based biosensors show high sensitivity and selectivity. However, they have few shortcomings, like complicated and costly fabrication process, degradation of the enzyme in biological environments after longer use and sometimes low sensitivity owing to indirect electron transfer. To overcome the disadvantages of enzyme-based glucose biosensors, the development of non-enzymatic glucose biosensors is highly desirable. Many experiments have been performed to develop non-enzymatic glucose biosensors. Pure and copper doped ZnO nanoparticles have been investigated for their use in non-enzymatic glucose sensors.[15] Both can be used as a cost effective non-enzymatic glucose biosensor. However, Cu doping improves the sensitivity for glucose detection. In another study, the detection limit of ~0.4 μM has been achieved by the glucose sensor in a vertically grown ZnO nanorods on fluorine doped tin oxide (FTO) electrode after modifying the ZnO surface with CuO nanoparticles.[16] It is reported that zinc oxide (ZnO) nanorod powder surface-coated

with carbon material, which has low detection limit of 1 mM with a linear range from 0.1 mM to 10 mM, can be applied as cost-effective non-enzymatic glucose biosensors.[17]

The development of non-enzymatic biosensors needs a detailed understanding of interfacial chemistry at the nano-bio interface. The interactions of biomolecules with metal oxide surfaces in aqueous solutions is dominated by adhesion of biomolecules on to the surfaces via water molecules. The interactions of amino acids with ZnO surfaces have been studied using molecular dynamics simulations.[18] It is observed that the first layer of water molecules near the surfaces affects the approach of the amino acids to the surface. Another MD simulations study on amino acid side chain analogous and a titanium binding peptide on TiO2 (100) surface shows that polar and aromatic side chain are the strongest binders to TiO2 surface.[19]

Here we investigate interactions between glucose molecules and common surfaces of ZnO, $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ in hydrated conditions, using the molecular dynamics (MD) simulations. We further compute the potential of mean force (PMF) using the umbrella sampling technique to quantify energetics of surface specific interactions. These calculations show that glucose exhibits strongest adsorption on $(10\bar{1}0)$ surface. We confirm the stability of glucose on this surface also by quantum mechanical (QM) calculations based on density functional theory (DFT) method.

5.2 Methods

5.2.1 Quantum mechanical (QM) calculations

All QM calculations are performed within density functional theory (DFT) (see Appendix I, Chapter 4 for theoretical background), using the projector augmented-wave method[20,21] as implemented in the Vienna Ab initio Simulation Package (VASP)[22]. We use generalized gradient approximation (GGA) with the Perdew, Burke, and Ernzerhof (PBE) correlation

functional. We use a plane wave cut off energy of 500 eV, and the total energies are converged to within 10^{-4} eV/atom. The optimization of ZnO bulk is done using a k-point mesh of 8x8x4 making sure that the Γ point is included in the mesh.

A supercell of $2 \times 3 \times 3$ is made using the optimized structure of ZnO bulk. We construct slabs with desired exposed ZnO surface from the supercell and modelled by including a vacuum of 15 Å above crystallographic direction. Self-consistent calculations of the surfaces are done using a k mesh of $1 \times 4 \times 4$ for $(10\bar{1}0)$, $4 \times 1 \times 4$ for $(11\bar{2}0)$ and $4 \times 4 \times 1$ for the slab containing both $(000\bar{1})$ and (0001) surfaces respectively. We also optimize glucose molecule by keeping it inside a box of dimension $2 \times 2 \times 2$ nm^3. We use a k mesh of $8 \times 8 \times 8$. The optimizations of the surfaces are done by keeping the middle two layers fixed, while the other layers of both sides are relaxed in all directions.

We also consider $(10\bar{1}0)$ surface with glucose and water for QM calculations. We perform single point calculations of the systems using a k mesh of $1 \times 4 \times 4$ considering snapshots from the molecular dynamics trajectory. The total number of atoms of the surfaces are restricted to 192 (96 Zn and 96 O atoms) for making the systems computationally affordable. We calculate the electronic charge for each atom using the calculated charge density of the structure using a grid-based Bader charge analysis.[30,31] The formula for calculating the charge transfer between any two species (*A* and *B*) is,

$$\Delta q = -e\Delta n = n_{atom} - n_{valence},$$

where the number of electrons calculated around an atom is n_{atom} and $n_{valence}$ is the number of valence electrons considered for the corresponding atom.

5.2.2 Molecular dynamics (MD) simulations

Next, we prepare the systems for the MD simulations (see Appendix I, Chapter 2 for theoretical background) as detailed below. First, we prepare a system where we add 100 glucose molecules in water in a box of dimension $6.89 \times 4.95 \times 5.29$ nm^3. The system contains a total number of 20970 atoms with the volume of the box 180.55 nm^3. We prepare the system of ZnO surface with water as follows: First, the optimized structures of the slabs are extended into two periodic directions in a fashion $1 \times 5 \times 5$ for $(10\bar{1}0)$ surface, $5 \times 1 \times 5$ for $(11\bar{2}0)$ surface and $5 \times 5 \times 1$ for the slab containing both (0001) and $(000\bar{1})$ surfaces. The total number of atoms of each extended slab is 2400 (1200 Zn and 1200 O). The surfaces $(10\bar{1}0)$ and $(11\bar{2}0)$ are placed in the middle of the box of dimensions $6.89 \times 4.95 \times 5.29$ nm^3 and $5.69 \times 6.0 \times 5.29$ nm^3 respectively. Similarly, the slab, containing both (0001) and $(000\bar{1})$ surfaces is placed in the middle of the box of dimension $5.69 \times 4.94 \times 6.43$ nm^3. Subsequently, the boxes are solvated with 6190 water molecules.

We take an equilibrated system of ZnO slab with water. We remove few water molecules and add 100 glucose molecules into the system for each cases. The total box volume (180.55 nm^3) and the total number of atoms (20970 atoms) for all the systems are same. As both sides of $(10\bar{1}0)$ and $(11\bar{2}0)$ surfaces are identical, all the analyses are averaged over both sides of the surfaces.

The force field of gromos54a7[23] and spc/e water model are used in this study. The Lennard-Jones (L-J) parameters for ZnO are calculated through density functional theory[24]. The L-J parameters for Zn are σ_{Zn}= 0.171 nm and ε_{Zn}= 1.254 kJ/mol, and for O, σ_O= 0.213 nm and ε_O= 0.418 kJ/mol. The parameters corresponding to the atom Zn interacting with atom O are determined through $\sigma_{ZnO} = \frac{1}{2}(\sigma_{Zn} + \sigma_O)$ and $\varepsilon_{ZnO} = \sqrt{\varepsilon_{Zn}\varepsilon_O}$. For each surface, both Zn and O atoms are restrained by a harmonic potential of force constant 10000 kJ mol^{-1} nm^{-2},

so that during the simulation the volume of the box does not fluctuate widely. Particle mesh Ewald summation method for long ranged electrostatic interaction is used with 0.1 nm grid spacing and 10^{-6} convergence criterion. We use a cut off of 1 nm for both L-J and short ranged electrostatic interactions.

The steepest descent and conjugate gradient[25] algorithms are used to minimize the systems. Molecular dynamics simulations are performed using Gromacs 2019 program[26,27] at a temperature of 300 K and 1 atm pressure in a canonical (NVT) ensemble starting from the energy minimized structure. Modified Berendsen thermostat is used to maintain constant temperature. A total simulation of 100 ns has been carried out for each system. All the analyses have been done from the equilibrated trajectories of last 50 ns.

5.2.3 Calculations of density

The effective density is calculated from the formula, $\rho' = \frac{nM}{v'}$. Here, $v' = L_x \times L_y \times h'$, where L_x, L_y are the total length of the box in x and y directions, and h' is the effective height of the box. n is the number of glucose in which the centre of mass (COM) of glucose lie within volume $v'$of the simulation box, and M is the molar mass of glucose. When calculating the density of water, n becomes the number of water molecules in which O atoms of water molecules lie within v'. For each surface, we calculate the density of water both in presence and absence of glucose, and the density of glucose. We also compute the density of water around each glucose molecule in presence and absence of the surfaces. For the calculations of density of water both in presence and absence of glucose, and the density of glucose, we first, compute the mean of z coordinates of the atoms in the top layer of each surface to define the mean surface locations with respect to which the density is calculated as a function of separation in z, Δz. For the calculations of density of water both in presence and absence of

glucose, we take the distance between the COM of glucose and the O atoms of water molecules. A bin size of 0.2 Å is used.

5.2.4 Calculations of potential of mean force (PMF)

The adsorption free energy of glucose on different surfaces in aqueous solutions is calculated by using umbrella sampling method as implemented in Gromacs.[26,27] The details are given in Appendix I of this chapter. First, we consider a single glucose molecule, placed at a larger distance from the surface. Subsequently, the system is solvated with water. A molecular dynamics simulation of 10 ns has been carried out for equilibration, keeping the glucose fixed in its position. The last structure of this 10 ns run has been considered for steered molecular dynamics (SMD) simulations prior to umbrella sampling. In SMD simulations, the molecule is pulled towards the surface from its initial position, in z direction, Δz_{cm}, separating the COM of glucose and the top of the surface. The pulling is done through a dummy particle moving towards the surface with a constant speed of 1 nm ns^{-1} with the COM of glucose is dragged by a harmonic force of spring constant, k=1000 kJ mol^{-1} nm^{-2}. The conformations are saved within each of the intervals of 0.1 nm from the SMD simulation trajectory. In this way we have collected about 25 conformations for each of these surfaces. They are used for further run of 10 ns each (5 ns for equilibration), where the z location of the glucose is fixed and the COM moves within a sampling window of width Δz. Finally, PMF calculation is done by using weighted histogram analysis method (WHAM)[28] as implemented in g_wham analysis tool[29] in Gromacs.[26,27]

5.3 Results and discussion

5.3.1 Optimized structures of the slabs

The calculated lattice parameters for ZnO bulk system are a=3.28 Å and c = 5.29 Å, which are close to the experimental values.[32] The optimized geometries of different slabs and

that of glucose are shown in Figs. 1(a)-(d). In the top layer of the optimized structure of the slab with exposed $(10\bar{1}0)$ surface (Fig. 1(a)), Zn-O bonds are tilted with a relative displacement of 0.2 Å in the direction perpendicular to the surface between where O is outward and Zn, inward. The Zn-O bond lengths decrease by 0.14 Å with respect to the bulk. Our results agree with the previous theoretical studies.[33,34] For slab with exposed $(11\bar{2}0)$ surface (Fig. 1(b)), the Zn atoms in the top layer show relative inward displacement of 0.2 Å in the direction perpendicular to the surface. However, O atoms remain almost same as in their bulk positions in agreement with previous theoretical works.[33,35] The $(11\bar{2}0)$ surface shows a groove like structure. The contraction of Zn-O bond length on the top surface is small (0.09 Å). The top and bottom layers for both $(10\bar{1}0)$ and $(11\bar{2}0)$ surfaces are identical. The top layer of the slab with exposed (0001) surface is Zn terminated, whereas the bottom exposed surface $(000\bar{1})$ is O terminated (Fig. 1(c)). As a result, a net dipole moment arises in the slab, rendering the slab polar. We also show the optimized structure of glucose in Fig. 1(d).

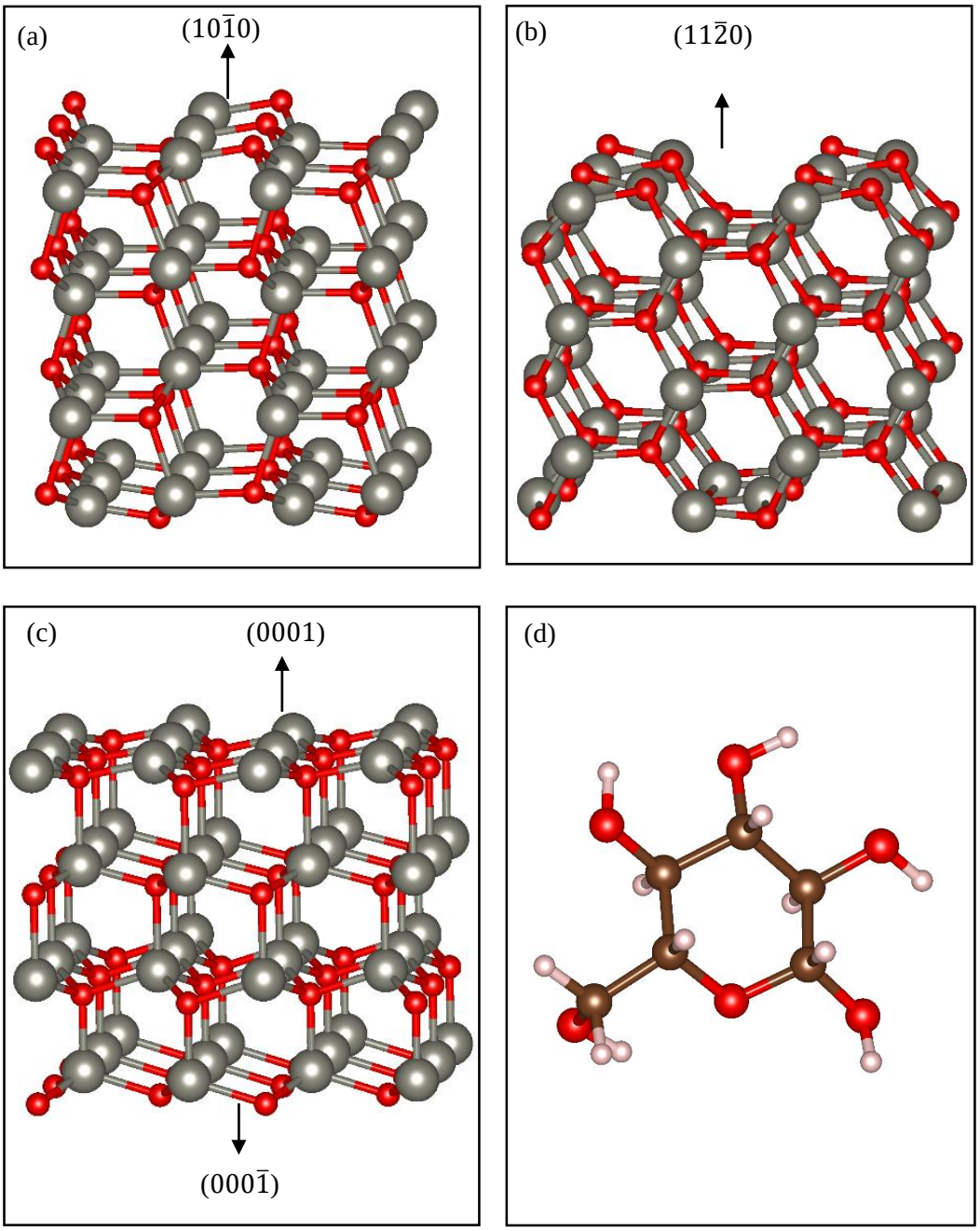

Fig. 1. Optimized geometries of (a) $(10\bar{1}0)$, (b) $(11\bar{2}0)$, (c) (0001) and $(000\bar{1})$ surfaces, and (d) glucose.

We carry out MD simulations in different situations: (a) glucose in water without surface; (b) ZnO surface with water; (c) ZnO surface with water and glucose. The schematic diagrams of the simulated systems for one such surface are shown in Fig. 2(a), (b) and (c) respectively. First, we illustrate the hydration of the chemical moieties and then describe the adsorption of glucose to the hydrated surface.

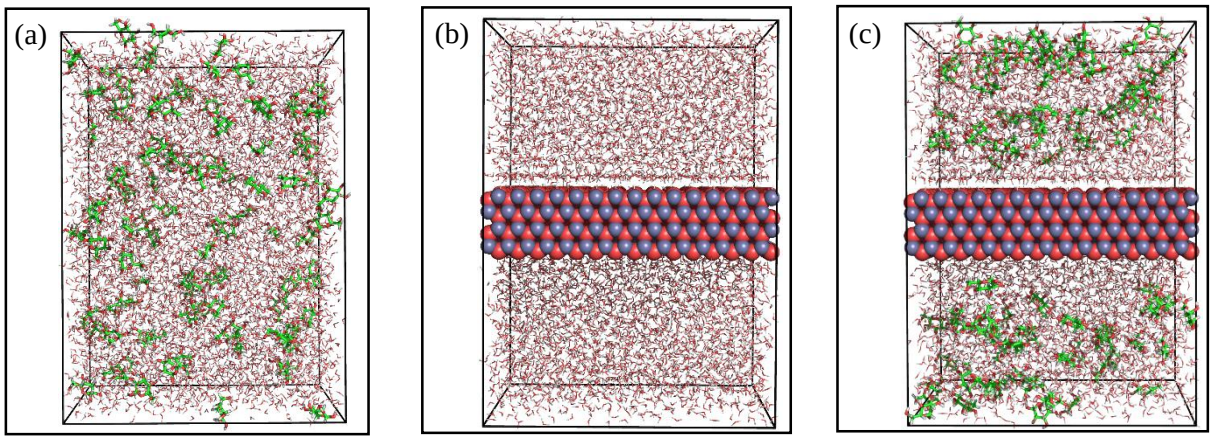

Fig. 2. Schematic diagrams of (a) glucose with water, (b) ZnO surface with water and (c) ZnO surface with glucose and water.

5.3.2 Hydration of glucose

We carry out simulation of glucose molecules in water alone without any surface. The density of water ρ_{wg} (r) around glucose molecules are shown in Fig. 3(a). Here, r is the distance between the COM of glucose and the COM of water molecule. The density consists of a peak at 0.47 nm with height of 1.357 gm/cc. The density finally saturates to bulk water density (1 gm/cc) for large enough r. We calculate hydrogen bonds (HBs), n_{HB}^{g}, between glucose and water molecules. In order to calculate HBs, we consider the following protocol: (a) the distance between the donor O and the acceptor O is less or equal to 0.34 nm and (b) the O-H (donor) – O (acceptor) angle is less or equal to 150^0. We observe that n_{HB}^{g}, shown in Fig. 3(b) is approximately 4 per glucose molecule.

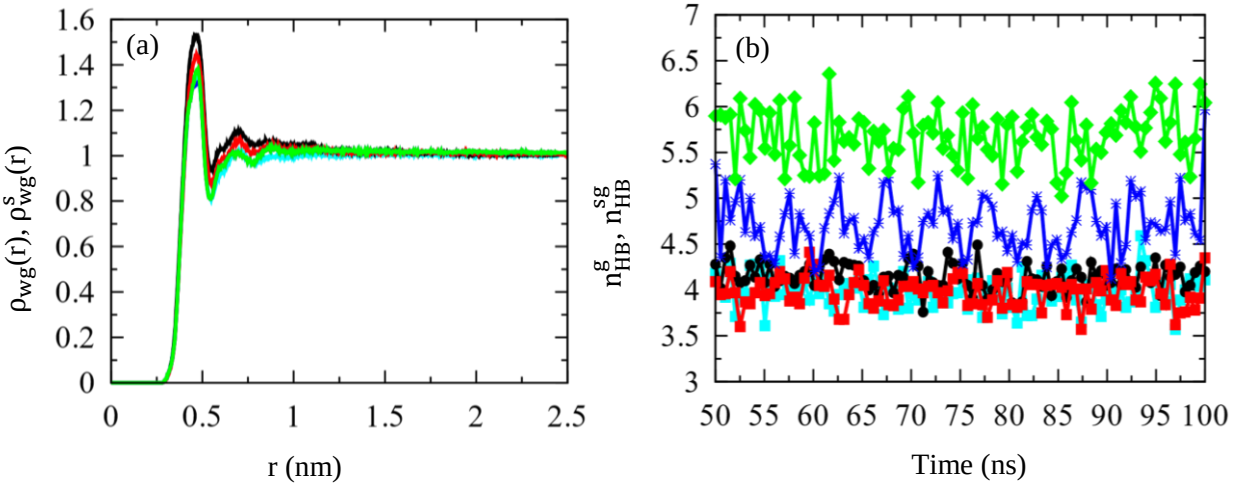

Fig. 3. (a) Density of water around glucose molecules in absence ($\rho_{wg}(r)$) and in presence ($\rho^{s}_{wg}(r)$) of the surfaces. (b) Fluctuations of number of hydrogen bonds (HBs) (n^{g}_{HB}, n^{sg}_{HB}) between glucose and water with time. The colors in black, red, blue and green represent the results for $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ surfaces. The results in absence of the surfaces are illustrated in cyan.

5.3.3 Hydration of ZnO surfaces

Now, we consider ZnO surfaces with water only. We show density profile of water above the surfaces, illustrated in Fig. 4. The water density profile ρ_{ws} (Δz) for $(10\bar{1}0)$ surface, shown in Fig. 4(a), consists of two well defined maxima at 0.12 nm (density of ~ 4.7 gm/cc) and 0.4 nm respectively. Water penetration occurs for $(11\bar{2}0)$ surface where first peak extends to $\Delta z<0$ in ρ_{ws} (Δz) (see Fig. 4(b)). This is due to grooves present in the surface and agrees with previous theoretical study.[18]. The strong density peak of 2.44 gm/cc at 0.178 nm has been found for this surface. ρ_{ws} (z) has three well defined maxima at 0.161 nm, 0.42 nm and 0.61 nm are observed for (0001) surface with maximum density 3.976 gm/cc at 0.161 nm, shown in Fig. 4(c). ρ_{ws} (z) in the vicinity of $(000\bar{1})$ surface (Fig. 4(d)) has two well defined maxima at 0.099 nm and 0.356 nm respectively. Thus, each slab has got strong hydration layer.

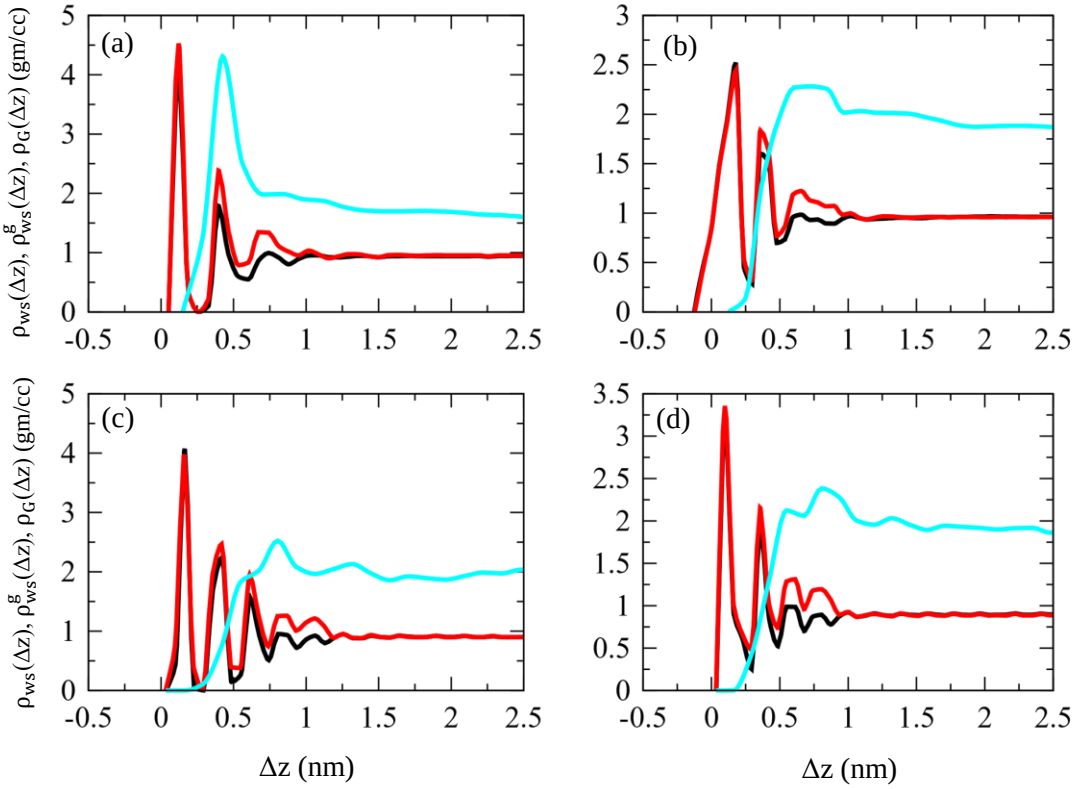

Fig. 4. $\rho_{ws}(\Delta z)$ (red), $\rho^{g}_{ws}(\Delta z)$ (black) and $\rho_G(\Delta z)$ (cyan) data for (a) $(10\bar{1}0)$, (b) $(11\bar{2}0)$, (c) (0001) and (d) $(000\bar{1})$ surfaces. The data for glucose density are multiplied by a factor of 9.

We capture the microscopic details of the water arrangement in the hydration layer. To facilitate the discussion, we show snapshots of water molecules above the surfaces of different slabs in Fig. 5. Three types of orientations of water molecules above the $(10\bar{1}0)$ surface can be seen in Fig. 5(a): (i) Water molecules where both the H atoms of each water molecule face to the surface O atoms, marked in blue solid ellipse; (ii) Water molecules with one H atom of each water molecule face to the surface O atoms of the top layer of the surface, while the other H face the O atoms in the second layer of the surface, as shown by the blue dotted ellipse; (iii) Water molecules where one H atom of each water molecule face to the surface O atom, while the other to the solvent, highlighted in yellow solid ellipse. The

orientations of most of the water molecules on $(11\bar{2}0)$ surface corresponds to the water molecules where one H atom of each water molecule face the surface O atom, while the other the solvent water molecule (see Fig. 5(b)). There are two types of orientations of water molecules found above the (0001) surface. In both the types, one H atom of each water molecule faces the surface O atom, and the other faces the solvent above the (0001) surface. They are arranged differently, shown in solid blue ellipse and dotted blue ellipse in Fig. 5(c). Similarly, two types of orientations of water molecules on $(000\bar{1})$ surface are observed. Here, both types of orientations belong to the water molecules where both H of each water molecule face the surface, but their arrangement are different, shown in Fig. 5(d). The orientations of the water molecules of the first layer above the different surfaces agrees with the previous theoretical results.[18]

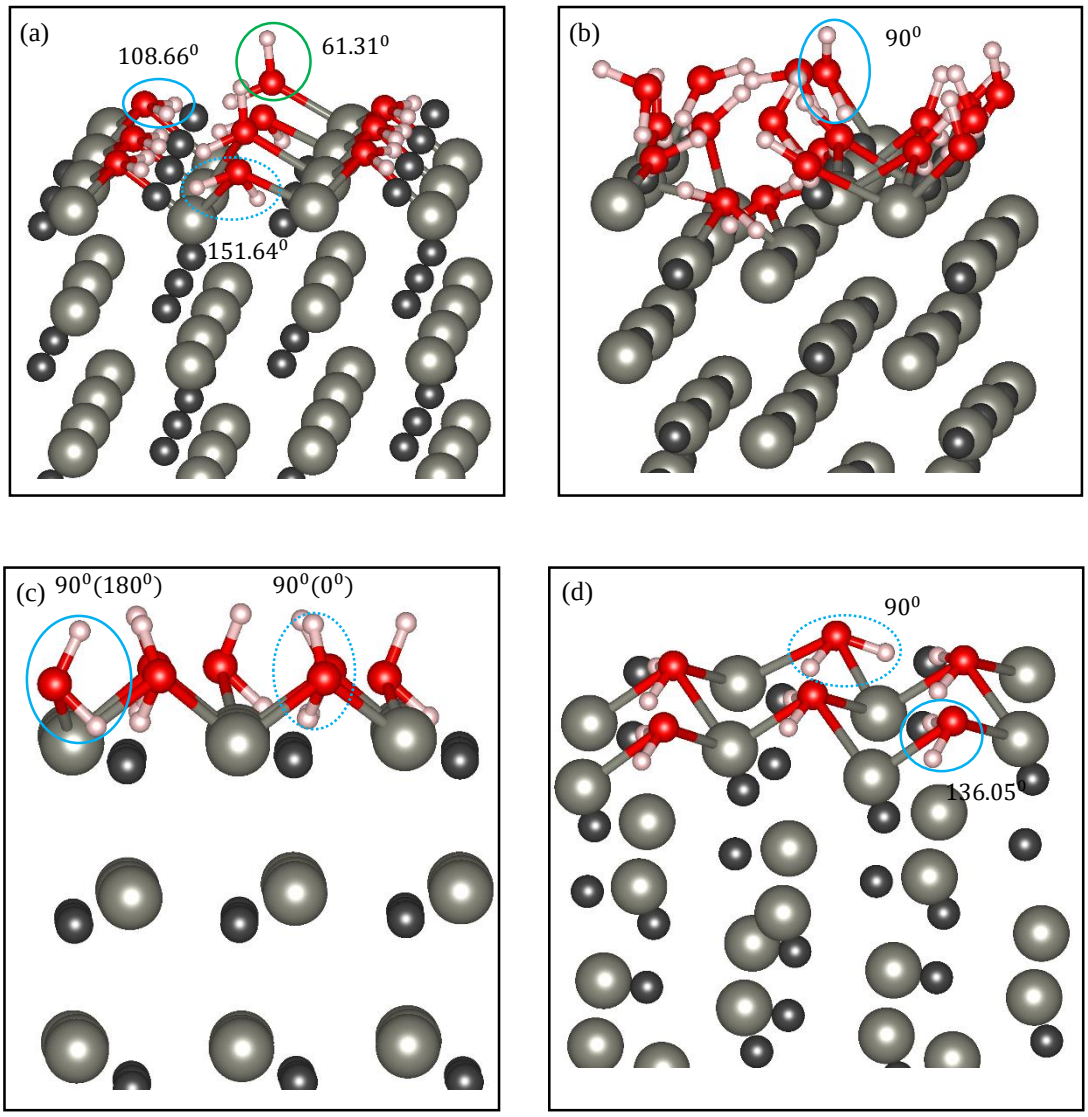

Fig. 5. Snapshots of solid-water interface for (a) $(10\bar{1}0)$, (b) $(11\bar{2}0)$, (c) (0001) and (d) $(000\bar{1})$ surfaces. Zn atoms are represented in grey, while O atoms in black. For water molecules, O atoms are illustrated in red, and H atoms are in white. The angle between the normal of the surface and the dipole moment of water for different orientations of water molecules are also shown. In (c), in addition to the angle between surface normal and dipole moment of water, the angle between tangent to the surface and the dipole moment is also calculated for different orientations of water molecules, shown in brackets.

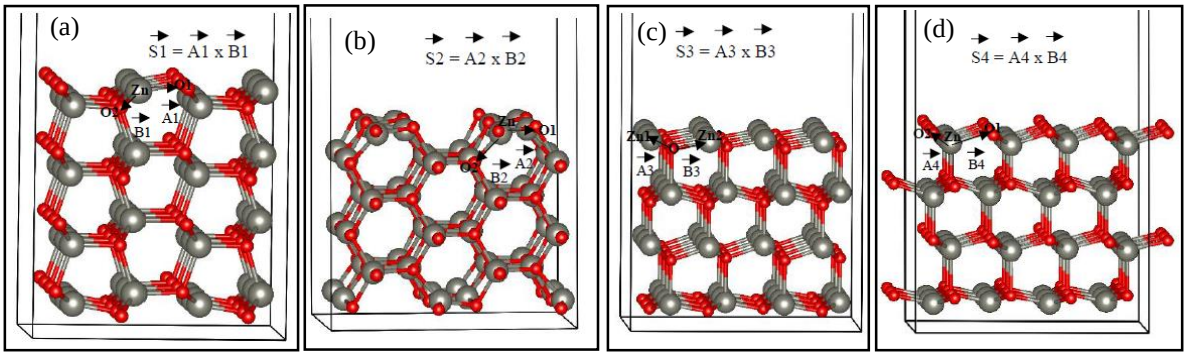

Fig. 6. Definition of surface vector for (a) $(10\bar{1}0)$, (b) $(11\bar{2}0)$, (c) (0001) and (d) $(000\bar{1})$ surfaces.

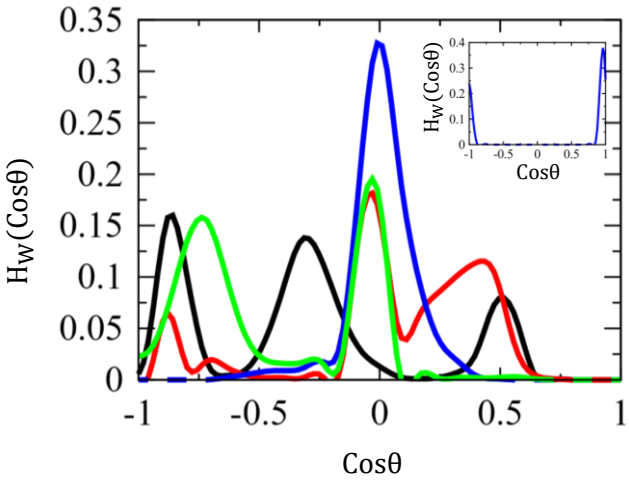

Fig. 7. $H_w(Cos\theta)$ of the orientations of the first layer of water molecules with respect to the surfaces. The orientations of water molecules with respect to the tangent of the (0001) surface is shown in inset. The results for $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ surfaces are illustrated in black, red, blue and green respectively.

We quantify the molecular orientation of water molecules by the orientation distributions $H_w(\cos\theta)$ of water molecules in the vicinity of the surface. Here θ is the angle between surface normal and the dipole vector of water. The surface normal is defined for $(10\bar{1}0)$ surface as follows: we consider two bonds, Zn-O1 and Zn-O2, in which a particular Zn atom in the top surface forms with two adjacent O atoms, shown in Fig. 6(a). The vectors correspond to Zn-O1 and Zn-O2 are $\overrightarrow{A1}$ and $\overrightarrow{B1}$ respectively. Now, $\overrightarrow{S1} = \overrightarrow{A1} \times \overrightarrow{B1}$. Similarly, we construct normals $\overrightarrow{S2}$, $\overrightarrow{S3}$ and $\overrightarrow{S4}$ for $(11\bar{2}0)$ (Figs. 6(b)), (0001) (Figs. 6(c)) and $(000\bar{1})$ (Figs. 6(d)) surfaces as well. The orientation profiles $H_w(\cos\theta)$ are shown in Fig. 7. Water molecules show three strong orientations on $(10\bar{1}0)$ surface at $\theta = 61.31^0$, 108.66^0 and

151.64^0 which are consistent to the orientations marked in Fig. 5(a). The non-polar surface $(11\bar{2}0)$, shows preferred orientation of water molecules at $\theta = 90^0$ in agreement to the snapshot in Fig. 5(b). The water molecules in the vicinity of the (0001) surface exhibits only one strong peak at $\theta = 90^0$. As both types of orientations of water molecules on this surface shows the same angle ($\theta = 90^0$) with respect to the normal of the surface, we also calculate the angle between the tangent of the surface with the water dipole moment. This gives the two orientations of water molecules at $\theta = 0^0$ and 180^0, shown in brackets in the snapshot of Fig. 5(c). The other polar surface $(000\bar{1})$ shows two maxima at $\theta = 90^0$ and 136.05^0 as in the snapshot of Fig. 5(d).

The number of hydrogen bonds, n_{HB} per surface O atom in absence of glucose for different surfaces with simulation time are shown in Fig. 8. It is observed that n_{HB} strongly depends on the type of the surface. The fluctuations in n_{HB} with simulation time are negligible for all these surfaces, indicating strong adsorption of the first layer of water to the surfaces. The average number of hydrogen bonds per O are 1.68 and 1.39 for $(10\bar{1}0)$ and $(11\bar{2}0)$ surfaces respectively. On (0001) surface, one of the H atoms of each water molecule face the solvent, while the other one to the O atom of the surface (see Fig. 5(c)). As a result, n_{HB} is 1, shown in the figure. The maximum number (n_{HB} = 2) is observed for $(000\bar{1})$ surface as for this surface both the H atoms of each water molecule are facing the surface and each surface O atom makes two HB with two water molecules, shown in Fig. 5(d).

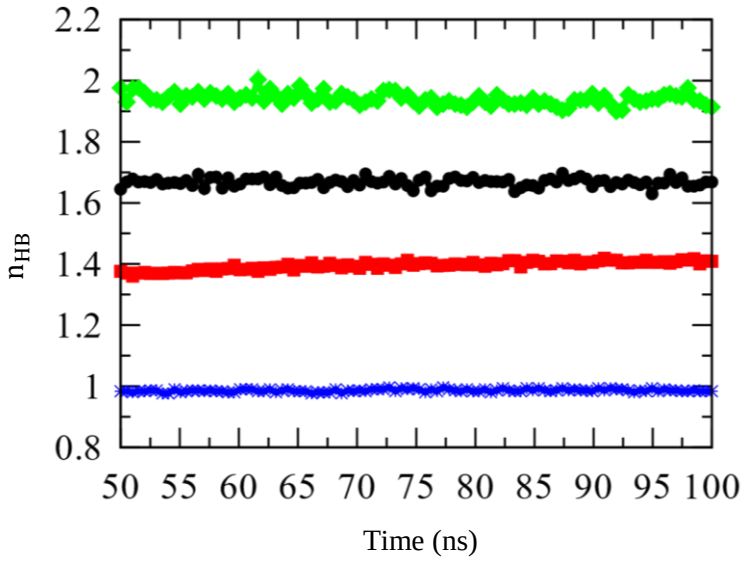

Fig. 8. The number of hydrogen bonds (n_{HB}) between the first layer of water molecules and $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ surfaces. The color in black, red, blue and green represent the results for $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ surfaces.

5.3.4 Glucose adsorption to hydrated ZnO surfaces

Next, we consider the system of ZnO along with water and glucose. The snapshots of glucose molecules above different surfaces are shown in Figs. 9(a)-(d). Most of the glucose molecules above $(10\bar{1}0)$ surface show flat conformations with respect to the surface (Fig. 9(a)). One such glucose molecule, having flat conformation is shown by the blue ellipse. For slabs with other exposed surfaces, glucose molecules are typically far away from the surface. Only few glucose come within 5 Å from the surface. Glucose molecules on $(11\bar{2}0)$ surface (Fig. 9(b)) do not show any specific orientations. The rings of the glucose molecules lie almost vertically with respect to the (0001) surface (Fig. 9(c)); one such glucose is marked in blue ellipse. For $(000\bar{1})$ surface (Fig. 9(d)), glucose molecules do not have any preferred orientations.

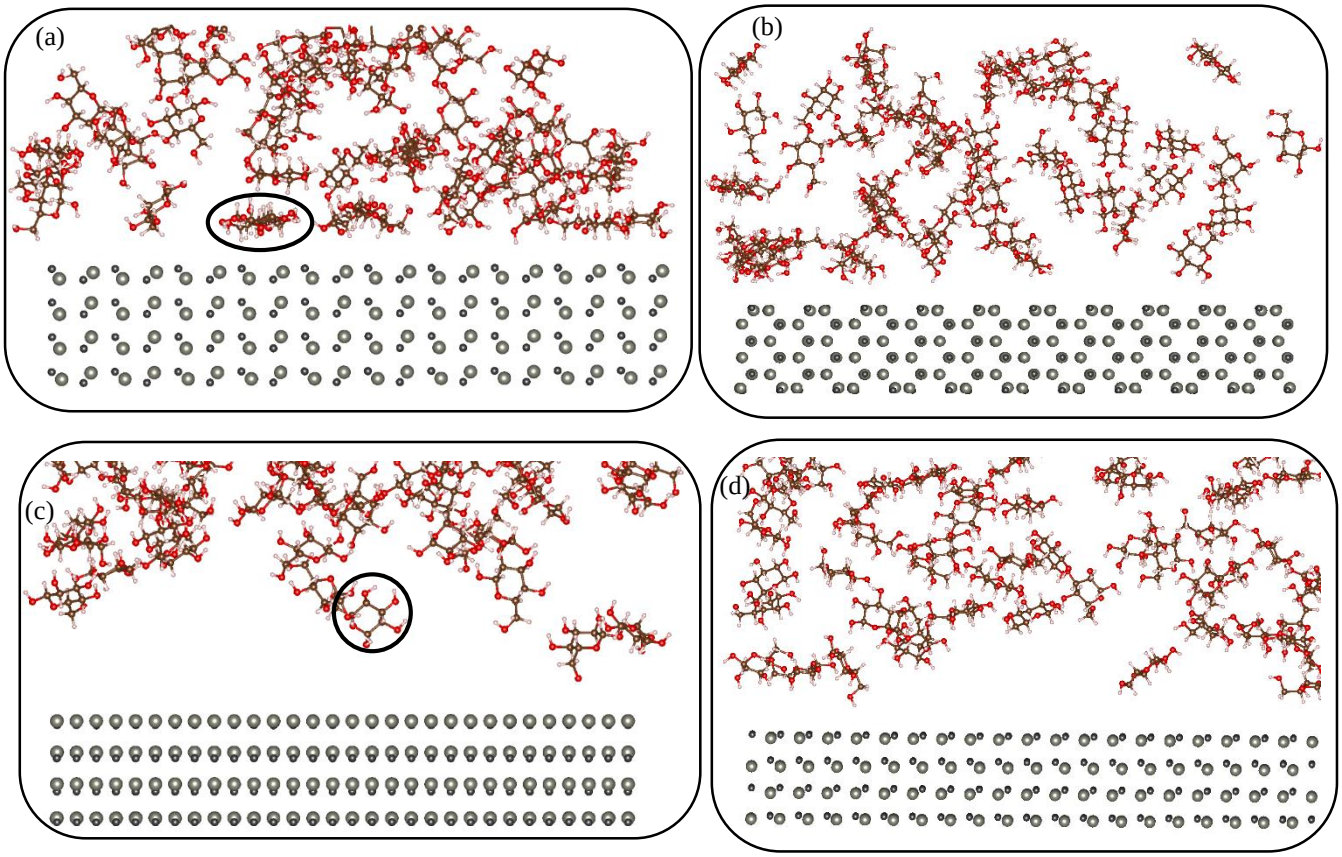

Fig. 9. Snapshots of glucose molecules above (a) $(10\bar{1}0)$, (b) $(11\bar{2}0)$, (c) (0001) and (d) $(000\bar{1})$ surfaces. Water molecules are not shown due to clarity. The most probable orientations of glucose molecules are highlighted in ellipse for $(10\bar{1}0)$ surface and in circle for (0001) surface.

Fig. 4 shows water density profile ρ_{ws}^{g} (Δz) in presence of glucose and glucose density profile ρ_G (Δz). We observe that the positions of the maxima of ρ_{ws}^{g} (Δz) do not change in presence of glucose compared to ρ_{ws} (Δz). The first peak heights are similar for both ρ_{ws}^{g} (Δz) and ρ_{ws} (Δz) for all surfaces. However, the other peak heights differ depending on the surface. The second peak height in ρ_{ws}^{g} (Δz) is reduced for $(10\bar{1}0)$ surface (Fig. 4(a)). ρ_G (Δz) appears approximately at the same positions (0.425 nm) as the second peak of water. This indicates that glucose molecules replace some water from the second layer to approach the surface. For the other surfaces, $(11\bar{2}0)$ (Fig. 4(b)), (0001) (Fig. 4(c)) and $(000\bar{1})$ (Fig. 4(d)), ρ_{ws}^{g} (Δz) are intact

compared to ρ_{ws} (Δz). ρ_G (Δz) do not have any significant peak in the density profile for these surfaces. Thus, glucose molecules are adsorbed on the $(10\bar{1}0)$ hydrated surface only.

The density of water around glucose molecules ρ^s_{wg} (r) in presence of the surfaces are shown in Fig. 3(a). All the distributions have similar peak values occurred at 0.47 nm, and then saturates to 1 gm/cc. The distributions in absence of surface are similar to those in presence of the surfaces. Thus, glucose molecules interact with the surfaces in hydrated conditions. We calculate the number of hydrogen bonds between glucose and water molecules, per glucose molecule in presence of surfaces with simulation time, shown in Fig. 3(b). It is observed that the non-polar surfaces $(10\bar{1}0)$ and $(11\bar{2}0)$ show similar HB profile compared to the glucose free solution, while the average number of HBs for both the polar surfaces are larger compared to those for the non-polar surfaces and for glucose free solution. The $(000\bar{1})$ surface shows the highest number of average HB (5.75) per glucose molecule.

We calculate further the orientation distributions of glucose $H_G(\cos\varphi)$ with respect to ZnO surfaces where φ is the angle between the normal to the plane of the sugar ring and the surface. We consider those glucose molecules for which any atom of glucose lies within 5Å from the top of the surfaces. We take two bonds of glucose between O atom and two adjacent C atoms in the ring, i.e., Og-C1 and Og-C2 bonds, labelled as $\vec{E}$ and $\vec{F}$ vectors respectively in Fig. 10(a), so that the normal to the ring, $\vec{G} = \vec{E} \times \vec{F}$. Finally, we calculate the cosine of the angle φ between the surface normal ($\overrightarrow{S1}, \overrightarrow{S2}$, $\overrightarrow{S3}$ and $\overrightarrow{S4}$ respectively) and $\vec{G}$. The orientation profile of the ring of glucose with respect to the surface normal, $H_G(\cos\varphi)$ is shown in Fig. 10(b). $(10\bar{1}0)$ surface shows a strong orientation at $\varphi = 28.36^0$, while the other orientations are less favourable. This angle represents nearly a flat conformations of glucose with respect to the surface in Fig. 9(a). For $(11\bar{2}0)$ surface, glucose molecules do not have any preferred orientations. (0001) exhibits a strong orientation at $\varphi = 71.34^0$ in agreement to the snapshot

in Fig. 9(c). Glucose molecules on $(000\bar{1})$ surface do not show any preferred orientations (Fig. 10(b)).

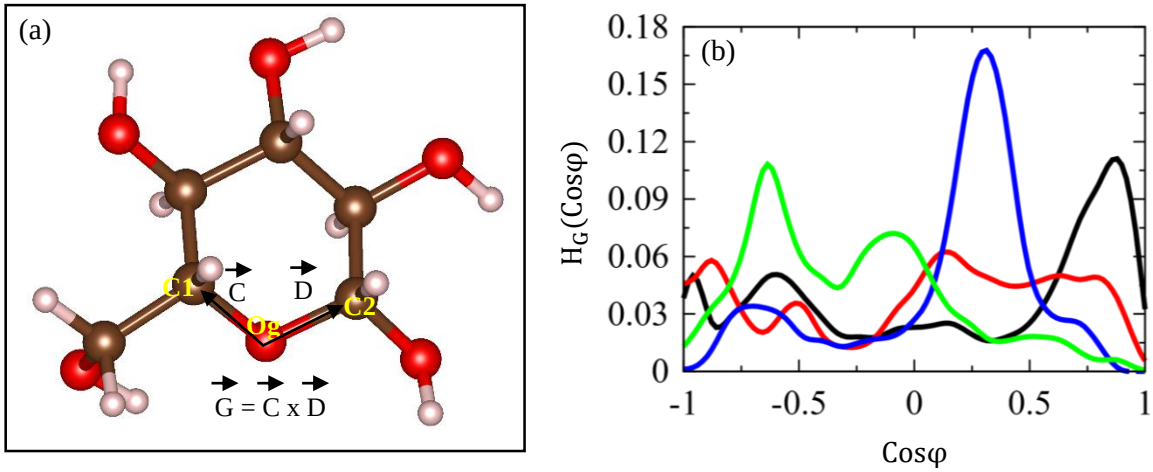

Fig. 10. (a) Calculation of the vector normal to the ring of glucose. (b) Orientation distributions, $H_g(Cos\varphi)$ of glucose with respect to the different surfaces. The colors in black, red, blue and green represent the results for $(10\bar{1}0)$, $(11\bar{2}0)$, (0001) and $(000\bar{1})$ surfaces respectively.

5.3.5 Potential of mean force (PMF)

The PMF profile, $V(\Delta z_{cm})$ as a function of Δz_{cm} where Δz_{cm} is the separation between the centre of mass (COM) of glucose and the mean surface for different surfaces are shown in Figs. 11(a)-(d). We observe multiple minima in $V(\Delta z_{cm})$ profile, one being global and the others the secondary minima. The global minimum in PMF profile for a particular surface corresponds to the adsorption free energy of glucose of that particular surface.

Two distinct well-defined minima in the PMF profile are observed for $(10\bar{1}0)$ surface in Fig. 11(a). The global minimum occurs at 0.429 nm with adsorption free energy -3.814 kJ/mol and a secondary minimum at 0.714 nm with a potential depth -1.114 kJ/mol. The barrier height between secondary and the global minima is 0.628 kJ/mol smaller than the room temperature thermal energy, 2.479 kJ/mol. So, when a glucose molecule approaches the surface, it enters the secondary minimum, but easily escape from the secondary minimum, and enter the global

minimum. Inside the global minimum, glucose is trapped and unable to escape from the global minimum since the barrier height (4.6 kJ/mol) has now surpassed that of room temperature energy. The value of Δz_{cm}, corresponding to the binding energy value agrees with the position (0.425 nm) of the peak where maximum density of glucose occurs (See Fig. 4(a)).

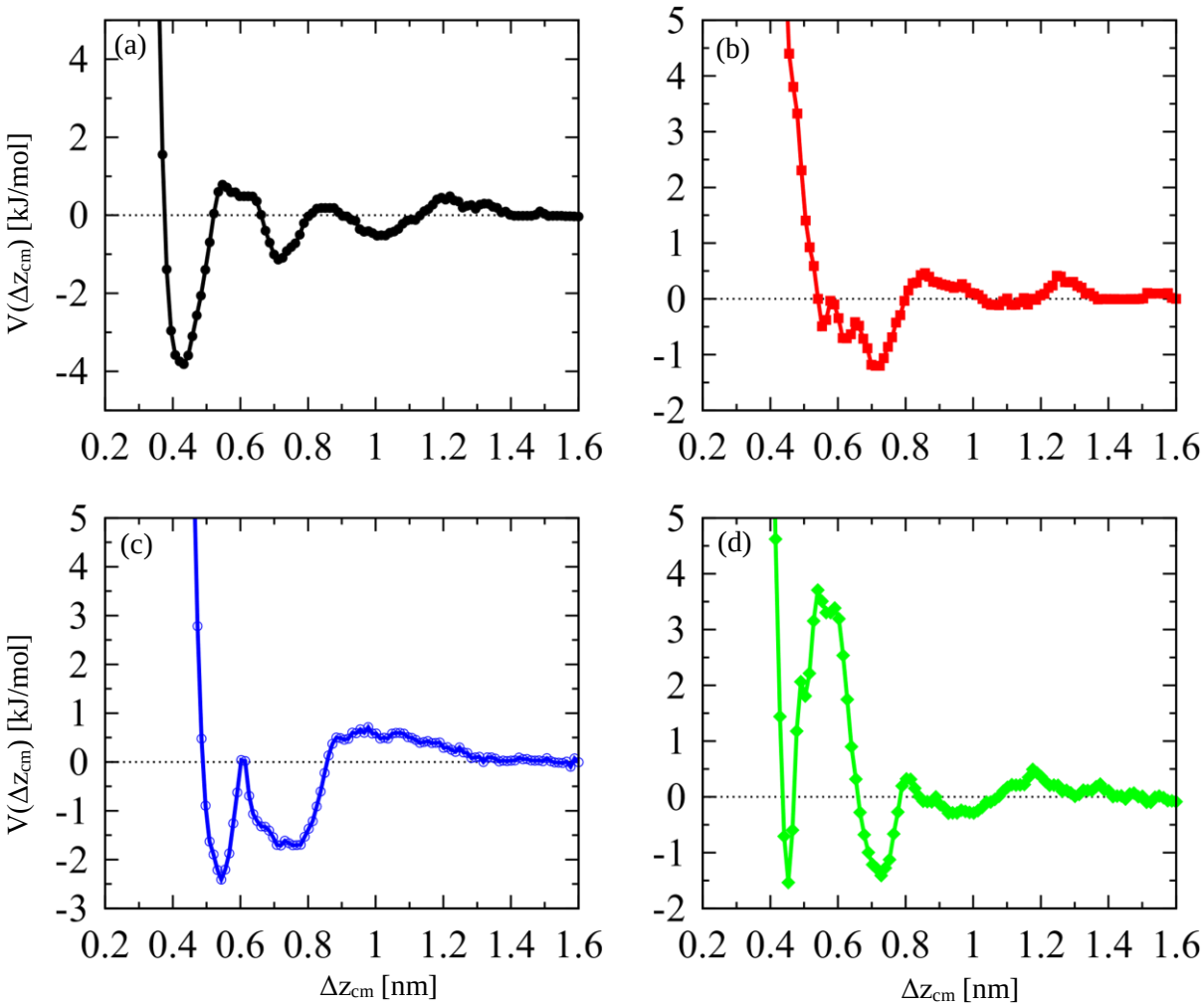

Fig. 11. Potential of mean force (PMF) profiles for glucose adsorption on (a) $(10\bar{1}0)$, (b) $(11\bar{2}0)$, (c) (0001) and (d) $(000\bar{1})$ surfaces respectively.

The $(11\bar{2}0)$ surface shows a global minimum at 0.714 nm with adsorption free energy value of -1.2 kJ/mol, shown in Fig. 11(b). The secondary minima occur at 0.554 nm and 0.624

nm with potential depths -0.504 kJ/mol and -0.704 kJ/mol respectively. The barrier height between any two minima is smaller than the room temperature energy. So, glucose can make a transition between any two minima. As a result, the adsorption of glucose on this surface is weak.

The polar surface (0001) shows global minimum at 0.544 nm with adsorption free energy -2.406 kJ/mol, shown in Fig. 11(c). The other two secondary minima for this surface occur at 0.717 nm and 0.769 nm with same potential depths -1.706 kJ/mol separated by a negligible barrier height. Hence, the barrier height between any of the two secondary minima and the global minimum is -1.706 kJ/mol which is smaller than the room temperature energy. So, glucose can easily make transitions to among the minima. As a result, there is no strong adsorption.

For $(000\bar{1})$ surface, there are two minima occur at 0.45 nm and 0.728 nm with similar well depth separated by a large barrier of 4.8 kJ/mol, shown in Fig. 11(d). The adsorption free energy value is -1.49 kJ/mol occurred at 0.45 nm. The potential depth of the other minimum, occurred at 0.728 nm is -1.39 kJ/mol. Hence, once the glucose molecule enters the secondary minimum, the transition between secondary to the global minima under room temperature is energetically unlikely.

The largest adsorption free energy found for $(10\bar{1}0)$ surface, happens due to two reasons: 1) glucose molecules come closer to this surface than they are to all other surfaces that increases the electrostatic interactions with the surface and 2) flat conformations of the glucose with respect to the surface, as we have already seen in the orientation profile of glucose on this surface (see Fig. 9(b)). When glucose is flat, it is more likely to have higher adsorption due to more contacts of glucose molecules with the water molecules that form bridges with the surface.

As $(11\bar{2}0)$ surface is rough and the water molecules have many orientations above this surface, glucose molecules are further away from this surface. But, for the (0001) surface, as the H atoms of each water molecule face the solvent, these H atoms form an almost frozen grid of positive partial charges (see Fig. 5(c)). So, the water molecules in the second layer are adjust in the way that the O atoms of each water molecule face the H atoms of the first layer. Hence, the second layer of water molecules for this surface is less mobile compared to those in the $(10\bar{1}0)$ surface. So, glucose molecules cannot penetrate even the second layer. Similar scenario holds for $(000\bar{1})$ surface, where both H atoms of each water molecule face the surface, while O atom to the solvent (see Fig. 5(d)). So, O atoms form a grid of partial negative charges, and the second layer of water molecules are adjusted in the way that H atoms of water molecules face the O atoms of the first layer. This makes the second layer less mobile. Hence, the breaking of the second layer by glucose molecules is unlikely.

The MD simulations results suggest that $(10\bar{1}0)$ surface has the largest adsorption free energy. As the water molecules form bridge between glucose molecules and the surface, we investigate the role of water molecules which form the bridge by quantum mechanical calculations in the framework of DFT. We choose two systems from the MD simulation trajectory with two different orientations of glucose molecules (system A, $\varphi = 51.08^0$ and system B, $\varphi = 36.40^0$), where both the orientations belong to the peak in the orientation profile of glucose with respect to the surface (see Fig. 10(b)). The systems are shown in Figs. 12(a) and (b). We perform single point calculations for both the systems.

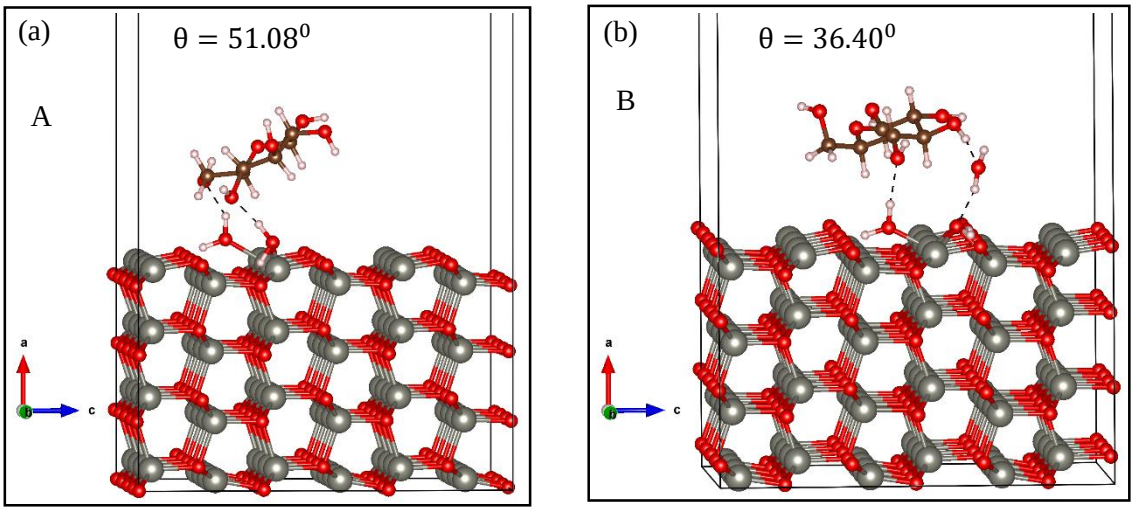

Fig. 12. QM structures, showing interactions between glucose and the surface via water molecules for (a) $\theta = 51.08^0$ and (b) $\theta = 36.40^0$.

The adsorption energy of glucose on hydrated ZnO surface is calculated by the following formula,

$$E_{ad} = E_{surf-glu-wat} - E_{surf} - E_{glu} - E_{wat},$$

where $E_{surf-glu-wat}$ is the total energy of the system containing ZnO surface, glucose and water, E_{surf} is the total energy of the ZnO surface, E_{glu} is the total energy of a single glucose molecule and E_{wat} is the total energy of a single water molecule. The adsorption energies of system A and system B are calculated to be -29.332 kJ/mol and -37.822 kJ/mol respectively. In system A, a charge of 0.02 e is transferred from water molecules to the glucose, while 0.04 e charge is transferred to the $(10\bar{1}0)$ surface from water molecules. This suggests that the water molecules help the glucose stabilize above the surface. However, the charge transfer is just too small to qualify for chemisorption.

5.4 Conclusions

We have investigated adsorption of glucose molecules to different surfaces of ZnO in aqueous solutions by means of MD simulations, coupled with umbrella sampling. The strong hydration layer on the surfaces hinders closer approach of the glucose molecules to the surfaces. However, considerable adsorption free energy of glucose on $(10\bar{1}0)$ surface indicates strong surface adhesion on the surface compared to the other surfaces. Hence, $(10\bar{1}0)$ surface is most preferable for adhesion of glucose in aqueous solution that may be useful in designing non-enzymatic glucose biosensors.

5.5 Appendix A

Details of the quantum mechanical calculations

A.1 PBE (Perdew-Burke-Ernzerhof) functional

To approximate the exchange-correlation energy, many different functionals are available. The details about the approximations to the exchange-correlation energy is already discussed in Appendix I of Chapter 4. Here, we have used PBE functional to approximate the exchange-correlation energy. PBE belongs to a class of GGA functional where no fitting parameters are used.

A.2 Pseudopotential

As the core and valence electrons' wavefunctions are orthogonal to each other, valence electrons show rapid oscillations near the core region. The size of the basis set would make it difficult to describe such functions using plane waves. One way to solve this issue is to use pseudopotentials. Pseudopotential is an approximation where the complicated effects of the motion of the core is replaced by an effective potential or pseudopotential. So, in this approach,

the core electrons are considered to be frozen, while the valence electrons are described explicitly. Projector augmented wave (PAW) method is a combination of pseudopotential and linear augmented-plane-wave (LAPW) method. This increases the computing efficiency of density functional theory calculations. LAPW method considers both valence and core electrons explicitly in DFT calculations, and it treats the full potential and charge density without any shape approximation.

A.2 Vienna Ab initio Simulation Package (VASP) code

The Vienna Ab initio Simulation Package (VASP) is a computer program that uses first-principles calculations for atomic scale materials modelling (e.g., electronic structure calculations). VASP uses density functional theory (DFT) or the Hartree-Fock (HF) approximation to construct an approximate solution to the many-body Schrödinger equation.

5.6 Appendix B

Details of the MD simulations

B.1 Umbrella sampling

For complex chemical and biological systems, Boltzmann sampling does not allow for proper exploration of phase space due to the fact that different volumes of phase space are separated by high energy barriers and the transitions between these volumes are very unlikely. To overcome the sampling problem, many techniques have been developed. 'Umbrella sampling or biased molecular dynamics' is one of the most used sampling methods for calculating free energy differences along a reaction coordinate. A reaction coordinate is a coordinate which represents the path or progress of a reaction. In umbrella sampling the reaction coordinate is divided into a number of windows. A target reaction coordinate is assigned to each window. The system is subjected to a bias potential $V_i(\xi)$, which is only a

function of the reaction coordinate, in each window. So, the biased energy of the system can be written as,

$$U_i^b(\mathbf{r}) = U_i(\mathbf{r}) + V_i(\xi) \quad (1)$$

However, the unbiased probability distribution is necessary to construct a free energy landscape along a reaction coordinate. The unbiased probability distributions ($P_i^u(\xi)$) can be written as,

$$P_i^u(\xi) = \frac{\int \exp[-\beta U_i(\mathbf{r})]\delta(\xi - \xi_i)d\mathbf{r}]}{\int \exp[-\beta U_i(\mathbf{r})]d\mathbf{r}]} \quad (2)$$

Whereas, the biased probability distributions along the reaction coordinate, $P_i^b(\xi)$ is given by

$$P_i^b(\xi) = \frac{\int \exp[-\beta U_i(\mathbf{r}) + V_i(\xi)]\delta(\xi - \xi_i)d\mathbf{r}]}{\int \exp[-\beta U_i(\mathbf{r}) + V_i(\xi)]d\mathbf{r}]} \quad (3)$$

therefore,

$$P_i^b(\xi) = \exp(-\beta V_i(\xi)) \frac{\int \exp[-\beta U_i(\mathbf{r})]\delta(\xi - \xi_i)d\mathbf{r}]}{\int \exp[-\beta U_i(\mathbf{r}) + V_i(\xi)]d\mathbf{r}]} \quad (4)$$

Equation 4 can be written in terms of equation 2 as follows,

$$P_i^u(\xi) = P_i^b(\xi) \exp[\beta V_i(\xi)]\langle \exp[-\beta V_i(\xi)]\rangle \quad (5)$$

So, the free energy can be obtained easily as follows,

$$A_i(\xi) = -\frac{1}{\beta}\ln\left(P_i^b(\xi)\right) - V_i(\xi) + F_i \quad (6)$$

where, $F_i = -\frac{1}{\beta}\ln(\langle \exp[-\beta V_i(\xi)]\rangle)$ depends on bias potential and is undetermined.

To obtain a global free energy curve, the free energy curves of each window are combined. Therefore, it is necessary to calculate F_i. The Weighted Histogram Analysis Method (WHAM), discussed in the next section, is a special technique for calculating potentials of mean force (PMFs) from a set of umbrella sampling simulations.

B.2 Weighted histogram analysis method (WHAM)

WHAM is used to obtain an unbiased probability distribution from biased probability distributions generated by umbrella sampling. A weighted average of the distributions of the individual windows yields the global distribution. The unbiased probability distribution is given by,

$$\mathrm{P^u}(\xi) = \sum_{i=1}^{N} w_i(\xi) P_i^u(\xi) \tag{7}$$

So, for a one dimensional reaction coordinate the WHAM equations are given by

$$\mathrm{P^u}(\xi) = \frac{\sum_{\mathrm{i}}^{\mathrm{N}} \mathrm{n_i}\, \mathrm{P_i^b}(\xi)}{\sum_{\mathrm{j}}^{\mathrm{N}} \mathrm{n_j} \exp[-\beta(\mathrm{V_j}(\xi) - \mathrm{F_j})]} \tag{8}$$

and

$$\exp\left[-\beta \mathrm{F_j}\right] = \int \exp[-\beta \mathrm{V_j}(\xi)] \mathrm{P^u}(\xi) \mathrm{d}\xi \tag{9}$$

where $\mathrm{n_i}$ is the total number of data points in $\mathrm{i^{th}}$ histogram.

To obtain an unbiased probability distribution, equations 8 and 9 are solved iteratively until the convergence is reached.

5.7 References

(1) Lynch, I.; Dawson, K. A. Protein-Nanoparticle Interactions. *Nano Today* **2008**, *3* (1–2), 40–47.

(2) Haynes, C. A.; Norde, W. Structures and Stabilities of Adsorbed Proteins. *J. Colloid Interface Sci.* **1995**, *169* (2), 313–328.

(3) Tawa, K.; Umetsu, M.; Nakazawa, H.; Hattori, T.; Kumagai, I. Application of 300× Enhanced Fluorescence on a Plasmonic Chip Modified with a Bispecific Antibody to a Sensitive Immunosensor. *ACS Appl. Mater. Interfaces* **2013**, *5* (17), 8628–8632.

(4) Ramani, M.; Nguyen, T. D. T.; Aryal, S.; Ghosh, K. C.; DeLong, R. K. Elucidating the RNA Nano–Bio Interface: Mechanisms of Anticancer Poly I: C RNA and Zinc Oxide Nanoparticle Interaction. *J. Phys. Chem. C* **2017**, *121* (29), 15702–15710.

(5) Sharma, V.; Shukla, R. K.; Saxena, N.; Parmar, D.; Das, M.; Dhawan, A. DNA Damaging Potential of Zinc Oxide Nanoparticles in Human Epidermal Cells. *Toxicol. Lett.* **2009**, *185* (3), 211–218.

(6) Rumyantseva, A.; Kostcheev, S.; Adam, P.-M.; Gaponenko, S. V.; Vaschenko, S. V.; Kulakovich, O. S.; Ramanenka, A. A.; Guzatov, D. V.; Korbutyak, D.; Dzhagan, V. Nonresonant Surface-Enhanced Raman Scattering of ZnO Quantum Dots with Au and Ag Nanoparticles. *ACS Nano* **2013**, *7* (4), 3420–3426.

(7) Gonçalves, R. A.; Toledo, R. P.; Joshi, N.; Berengue, O. M. Green Synthesis and Applications of ZnO and TiO2 Nanostructures. *Molecules* **2021**, *26* (8), 2236.

(8) Jafari, S.; Mahyad, B.; Hashemzadeh, H.; Janfaza, S.; Gholikhani, T.; Tayebi, L. Biomedical Applications of TiO2 Nanostructures: Recent Advances. *Int. J. Nanomedicine* **2020**, *15*, 3447.

(9) Jiang, J.; Pi, J.; Cai, J. The Advancing of Zinc Oxide Nanoparticles for Biomedical Applications. *Bioinorg. Chem. Appl.* **2018**, *2018*.

(10) Saniz, R.; Sarmadian, N.; Partoens, B.; Batuk, M.; Hadermann, J.; Marikutsa, A.; Rumyantseva, M.; Gaskov, A.; Lamoen, D. First-Principles Study of CO and OH Adsorption on In-Doped ZnO Surfaces. *J. Phys. Chem. Solids* **2019**, *132*, 172–181.

(11) Spencer, M. J.; Yarovsky, I. ZnO Nanostructures for Gas Sensing: Interaction of NO2, NO, O, and N with the ZnO (1010) Surface. *J. Phys. Chem. C* **2010**, *114* (24), 10881–10893.

(12) Gopikrishnan, R.; Zhang, K.; Ravichandran, P.; Baluchamy, S.; Ramesh, V.; Biradar, S.; Ramesh, P.; Pradhan, J.; Hall, J. C.; Pradhan, A. K. Synthesis, Characterization and Biocompatibility Studies of Zinc Oxide (ZnO) Nanorods for Biomedical Application. *Nano-Micro Lett.* **2010**, *2* (1), 31–36.

(13) Juska, V. B.; Pemble, M. E. A Critical Review of Electrochemical Glucose Sensing: Evolution of Biosensor Platforms Based on Advanced Nanosystems. *Sensors* **2020**, *20* (21), 6013.

(14) Rernglit, W.; Teanphonkrang, S.; Suginta, W.; Schulte, A. Amperometric Enzymatic Sensing of Glucose Using Porous Carbon Nanotube Films Soaked with Glucose Oxidase. *Microchim. Acta* **2019**, *186* (9), 1–7.

(15) Mahmoud, A.; Echabaane, M.; Omri, K.; El Mir, L.; Chaabane, R. B. Development of an Impedimetric Non Enzymatic Sensor Based on ZnO and Cu Doped ZnO Nanoparticles for the Detection of Glucose. *J. Alloys Compd.* **2019**, *786*, 960–968.

(16) Ahmad, R.; Tripathy, N.; Ahn, M.-S.; Bhat, K. S.; Mahmoudi, T.; Wang, Y.; Yoo, J.-Y.; Kwon, D.-W.; Yang, H.-Y.; Hahn, Y.-B. Highly Efficient Non-Enzymatic Glucose Sensor

Based on CuO Modified Vertically-Grown ZnO Nanorods on Electrode. *Sci. Rep.* **2017**, *7* (1), 1–10.

(17) Chung, R.-J.; Wang, A.-N.; Liao, Q.-L.; Chuang, K.-Y. Non-Enzymatic Glucose Sensor Composed of Carbon-Coated Nano-Zinc Oxide. *Nanomaterials* **2017**, *7* (2), 36.

(18) Nawrocki, G.; Cieplak, M. Amino Acids and Proteins at ZnO–Water Interfaces in Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* **2013**, *15* (32), 13628–13636.

(19) Brandt, E. G.; Lyubartsev, A. P. Molecular Dynamics Simulations of Adsorption of Amino Acid Side Chain Analogues and a Titanium Binding Peptide on the TiO2 (100) Surface. *J. Phys. Chem. C* **2015**, *119* (32), 18126–18139.

(20) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758.

(21) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 17953.

(22) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169.

(23) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40* (7), 843–856.

(24) Raymand, D.; Van Duin, A. C.; Baudin, M.; Hermansson, K. A Reactive Force Field (ReaxFF) for Zinc Oxide. *Surf. Sci.* **2008**, *602* (5), 1020–1031.

(25) Hestenes, M. R.; Stiefel, E. *Methods of Conjugate Gradients for Solving Linear Systems*; NBS Washington, DC, 1952; Vol. 49.

(26) Bekker, H.; Berendsen, H. J. C.; Dijkstra, E. J.; Achterop, S.; Vondrumen, R.; VANDERSPOEL, D.; Sijbers, A.; Keegstra, H.; Renardus, M. K. R. Gromacs-a Parallel Computer for Molecular-Dynamics Simulations. In *4th International Conference on Computational Physics (PC 92)*; World Scientific Publishing, 1993; pp 252–256.

(27) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1–3), 43–56.

(28) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021.

(29) Hub, J. S.; De Groot, B. L.; Van Der Spoel, D. G_wham□ A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* **2010**, *6* (12), 3713–3720.

(30) Tang, W.; Sanville, E.; Henkelman, G. A Grid-Based Bader Analysis Algorithm without Lattice Bias. *J. Phys. Condens. Matter* **2009**, *21* (8), 084204.

(31) Yu, M.; Trinkle, D. R. Accurate and Efficient Algorithm for Bader Charge Integration. *J. Chem. Phys.* **2011**, *134* (6), 064111.

(32) Özgür, Ü.; Alivov, Y. I.; Liu, C.; Teke, A.; Reshchikov, Ma.; Doğan, S.; Avrutin, V.; Cho, S.-J.; Morkoç, and H. A Comprehensive Review of ZnO Materials and Devices. *J. Appl. Phys.* **2005**, *98* (4), 11.

(33) Meyer, B.; Marx, D. Density-Functional Study of the Structure and Stability of ZnO Surfaces. *Phys. Rev. B* **2003**, *67* (3), 035403.

(34) Mora-Fonz, D.; Buckeridge, J.; Logsdail, A. J.; Scanlon, D. O.; Sokol, A. A.; Woodley, S.; Catlow, C. R. A. Morphological Features and Band Bending at Nonpolar Surfaces of ZnO. *J. Phys. Chem. C* **2015**, *119* (21), 11598–11611.

(35) Marana, N. L.; Longo, V. M.; Longo, E.; Martins, J. B. L.; Sambrano, J. R. Electronic and Structural Properties of the (1010) and (1120) ZnO Surfaces. *J. Phys. Chem. A* **2008**, *112* (38), 8958–8963.