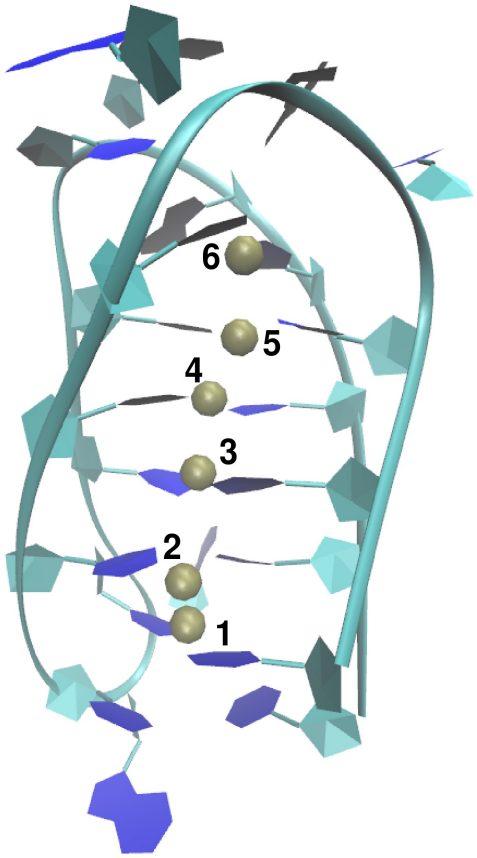
RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Deprotonation mechanism of a single-stranded DNA i-motif

Jens Smiatek and Andreas Heuer



An atomistic molecular dynamics simulation to study the unfolding and deprotonation mechanism of a single-stranded and fully protonated DNA i-motif.

Deprotonation mechanism of a single-stranded DNA i-motif†

Jens Smiatek,*[a] and Andreas Heuer[b]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

We present the detailed deprotonation mechanism of a fully protonated DNA i-motif as studied by all-atom Molecular Dynamics simulations. The associated unfolding of the i-motif is driven by the application of the Metadynamics method. The release of two protons, which stabilize the hemi-protonated cytosine pairs in the native i-motif, can be identified as the initial step for the occurrence of the unfolding process. By a systematic analysis of the proton motion, we are able to validate a two-step deprotonation mechanism. Our findings are in excellent agreement to experimental results that have been recently published.

In acidic solution around a pH value of 5, single stranded cytosine rich DNA can form specific non-helical structures, *e. g.* the i-motif which can be interpreted as a non-Watson-Crick like canonical conformation [1]. Since the discovery of the DNA i-motif in 1993 [1], a lot of effort has been spent to investigate this structure in theoretical as well as experimental work. It has been found that the preservation of the native structure is remarkably dependent on the presence of a slightly acidic solution. Hereby, hemi-protonated cytosine pairs (CC^+) can form which stabilize the i-motif in its typical form as shown in Fig. 1 [2–4]. In presence of a basic solution, deprotonation of the CC^+ pairs has been validated such that the native i-motif loses its stability and unfolds [5]. The systematic increase and decrease of the pH-value can be also used to reversibly switch between folded and unfolded configurations [4–7].

The biotechnological applications of this reversible folding mechanism include molecular nanomachines [5,8], switchable nanocontainers [9], sensors to determine the pH value inside living cells [10], building materials for logic gate devices [11] and detectors to distinguish between single walled and multi-walled carbon nanotube systems [12].

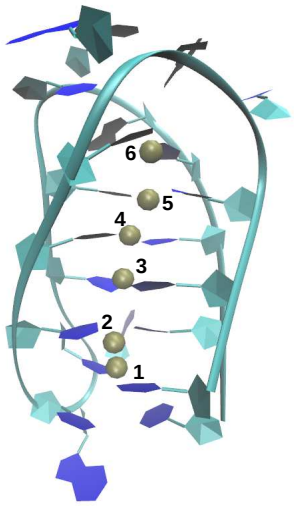


Fig. 1 Structure of the energy minimized DNA i-motif with the maximum number of six protons that form hemi-protonated CC^+ pairs.

Despite the technological usability, only few experimental and theoretical studies have thoroughly investigated the properties of i-motifs in terms of basic research like the occurrence of stable conformations and proton binding mechanisms. Combined experimental and computational studies have revealed that for basic pH values, the i-motif deprotonates and rapidly unfolds into hairpin configurations and partially unfolded structures [6,20–24]. Other numerical calculations instead have focused on the stability of hemi-protonated cytosine pairs in acidic solution [13–17]. The observed stability is specifically of importance due to the remarkable effect, that base-pair stacking interactions between multiple cytosine pairs alone have been found to be strongly repulsive [13]. An often used approach in computer simulations to study the stabilizing effects of protons without the application of pH dependent algorithms was given by the introduction of covalent bonds between hydrogen and nitrogen atoms which mimic hemi-protonated CC^+ pairs [14,17].

Further NMR studies in combination with Density Functional Theory calculations have investigated the potential en-

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

[a] *Institut für Computerphysik, Universität Stuttgart, 70569 Stuttgart, Germany. Fax: +49(0)711/68563658; Tel: +49(0)711/685-63757; E-mail: smiatek@icp.uni-stuttgart.de*
[b] *Institut für Physikalische Chemie, Universität Münster, 48149 Münster, Germany.*

ergy between the protonated cytosine pairs. It was found that the shape of the potential energy landscape for the coordinating protons can be sufficiently described by a double-well potential[25]. It has been also shown that an energy barrier of several kcal/mol at acidic pH values prevents the native i-motif structure from unfolding. Another interesting experimental result has revealed that the folding of i-motifs can be also driven by the presence of silver cations instead of an acidic pH value[26]. Although the authors have claimed that the resulting structure must not definitely coincide with the native i-motif configuration, the reversible folding and unfolding mechanism has been clearly demonstrated.

In summary, compared to the larger number of computational studies concerning the behavior of G-quadruplex structures, which is formed due to guanine-guanine base pairs (for an overview see Refs.[18,19]), detailed Molecular Dynamics simulations of the unfolding motion and the deprotonation mechanism of a DNA i-motif are rather scarce. An important experimental step towards a systematic understanding of these mechanisms has been recently performed by Liu and coworkers[6]. The detailed unfolding time for a 21-mer i-motif has been reported for a series of stopped flow circular dichroism (UV CD) experiments at different pH values. The corresponding reaction times have been determined to occur on a time scale between 0.1 s to 150 s (pH values between 8.22 and 6.52). The authors of the study were additionally able to validate the presence of a two step deprotonation process. Thus, for the first unfolding step, an initial release of two protons has been revealed which strongly determines the overall unfolding times.

In this article, we explicitly focus on the deprotonation mechanism and the corresponding unfolding process of a single stranded DNA i-motif in terms of biased Metadynamics/all-atom Molecular Dynamics simulations[29–31]. Our results reveal a two-step deprotonation mechanism which is initialized by the rate-limiting release step of two protons which is in excellent agreement to the aforementioned experimental results[6]. Furthermore we are able validate the occurrence of specific hairpin conformations which are closely related to previous studies on deprotonated DNA i-motifs[23,24,27]. The simulation details are in detail discussed in the supplementary material. We have modeled the protons as individual atoms but without the existence of covalent bonds with the nucleobases as it was proposed in Refs. [14,17].

The protonated DNA i-motif structure has been modeled to include six protons which mimic hemi-protonated cytosine pairs in agreement to the assumptions published in Ref.[28]. The energetically minimized structure is shown in Fig. 1. It can be clearly seen that the conformation is fully protonated with the maximum number of six coordinating protons in agreement to six hemi-protonated cytosine base pairs[28]. All simulations have been conducted by using the Metadynamics algorithm which helps to overcome energetic barriers and to accelerate rare events[29–31]. The corresponding pathways have been also shown to resemble the lowest free energy path[30,32,33] and are often in good agreement to the unbiased simulation trajectories[24,27].

After equilibration and initial release of the position restraints, the protons immediately (within 100 ps biased simulation time) change their positions which leads to a novel protonation state whose structure is shown in the supplementary material. It can be found that proton 4 as denoted in Fig. 1 in the rearranged conformation interacts with the oxygen atoms of two phosphate groups at different sides of the strand. We propose that the strong electrostatic repulsion between the protons as it is also discussed in the supplementary material can be considered as a reason for this behavior. The remaining five protons form a new protonation state where one proton coordinates roughly four nucleobases in an intercalating scheme. It has to be noted that the corresponding configuration is remarkably identical to G-Quadruplex-structures and intercalating potassium ions[19,34]. In fact, quantum chemical computations which take into account explicit solvent molecules would allow us to finally decide on the reliability of the derived configuration[35,36]. Several characteristics of oligonucleotides like the larger number of considered particles, severe problems with the underlying electronic density functionals[37], or the prohibitively long time scales for the proton dynamics strongly reduce the computational feasibility of a systematic QM study. Despite this remark, a recent publication[26] has also pointed out the existence of a diminished protonation state in presence of silver cations due to the presence of strong electrostatic repulsions in addition to large atom sizes which has been also thoroughly discussed in Ref.[35].

In order to understand the dynamic unfolding behavior of the DNA i-motif and the associated influence of the protons, we have monitored the actual number of contact pairs within a maximum distance of 0.6 nm between the protons and DNA atoms N_P normalized to the initial structure. The averaged results over ten independent Metadynamics simulation runs are shown in Fig. 2. It can be clearly seen that the aforementioned rearrangement of proton 4 is significantly pronounced in the contact pair representation after an interval of 250 ps. The corresponding rearrangement of the other protons during the first nanosecond can be recognized by pronounced fluctuations in Fig. 2. The resulting configuration for the remaining five protons leads to higher contact pair numbers for protons 3,5 and 6 which is indicated by $N_P \geq 1$. Contrarily, protons 1 and 2 do not show this increased contact number instead of a nearly monotonous decrease for proton 2 and fluctuating values for proton 1. The slightly instable position of proton 1 due to a smaller number of contact pairs as it can be seen in Fig. 1 be-

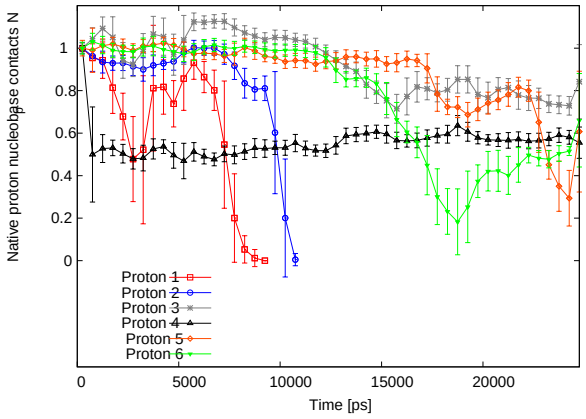


Fig. 2 Number of contact pairs N_P between protons and atoms of the DNA i-motif. A pair is present if the maximum distance of 0.6 nm between the atoms is not exceeded. The early decrease in the number of contact pairs for proton 4 due to rearrangement effects with the backbone is evident. The values are averaged over ten independent biased Metadynamics simulations.

comes also evident by these strong fluctuations. Finally after 8 nanoseconds, proton 1 is released and proton 2 follows after roughly 12 nanoseconds. Hence, a nearly concerted release for these two protons can be proposed. With regard to the complete deprotonation mechanism of the i-motif, our results clearly reveal a two step process with the initial release of two protons in good agreement to recent experimental findings[6].

This assumption is furthermore supported by the calculation of the Pearson correlation coefficients for individual protons and their corresponding number of proton-nucleobase atom contact pairs. The calculations are presented in the supplementary material. We have found a strong agreement between the motion of protons 1 and 2 which is significantly larger compared to the other proton pairs except protons 5 and 6. We are therefore confident that these results indicate a concerted release to follow a combined dynamic behavior.

In terms of the simulation times, it has to be noticed that the presence of biasing potentials as given by Metadynamics accelerates the sampling of rare events such that the observed simulation times cannot be mapped to real time units[31]. However, due to recent results where the trajectories of biased simulations are compared to unbiased simulations[24], we are confident that the observed biased deprotonation mechanism resembles the unbiased mechanism in basic solution in good agreement.

Furthermore we have detected stable configurations of the DNA i-motif along the pathways by the evaluation of the free energy landscapes as it is discussed in the supplementary material. The stable structures are shown in Fig. 3. It can be shown that conformation I (*native i-motif*) and II (*hairpin structure*) correspond to global free energy minima. The additional configuration III represents a less stable expanded state. The reason for the stability of the hairpin structures was in detail discussed in Ref.[23] where it was found that this observation is strongly related to a significant ordering of water molecules. It is evident that four protons remain attached to the hairpin configuration. The further unfolding into random coil swollen configurations (III) does not merely include a fixed number of released protons. Thus we can conclude that the proton release step is associated to the transition from configuration (I)→(II). The further unfolding corresponds to the exploration of the phase space in terms of local free energy minima. By having a look at the free energy landscapes in the supplementary material, it becomes obvious that $k_{12} \ll k_{23}$ and therefore the transition from the native structure to the hairpin conformation can be interpreted as the rate determining step for the overall unfolding mechanism. Hence, all our results are in excellent agreement to recent experimental findings[6]. Although one has to be careful due to severe limitations of the model and the nucleic acid force fields for long simulation times[19], we are confident that our approach allows a first computational insight into the unfolding mechanism of DNA i-motifs.

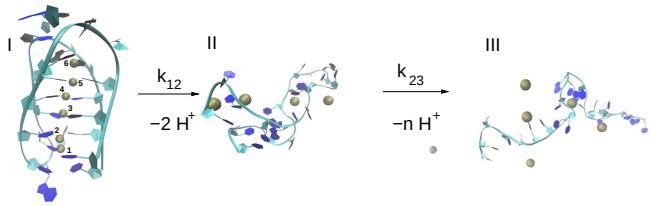


Fig. 3 Stable conformations of the DNA i-motif with rate constants and number of released protons. Structure I is the native i-motif, structure II illustrates hairpin configurations whereas structure III represents a swollen state.

In summary, we have found that after the release of two protons, the DNA unfolds into hairpin configurations as well as fully unfolded structures which is in agreement to recent experimental studies of deprotonated DNA i-motifs[23,24]. In the following step, the hairpin conformation unfolds into random coil swollen configurations where additional proton-DNA contacts have not necessarily to be broken which means that the protons remain attached to the DNA strand. Hence, we can assume that the reported mechanism resembles the observed experimental facts[4,6,20–24] and can be interpreted as a further step towards the understanding of the DNA i-motif unfolding motion. We therefore hope that our results will help to understand the fundamental dynamics of the DNA i-motif in more detail.

We have strongly benefitted from discussions with Dongsheng Liu, Maria Fyta, Jozica Dolenc, Malte Drescher, Jörg Harting and Vasileios Tatsis and thank the Deutsche Forschungsgemeinschaft through the transregional collaborative research center TRR 61 for financial funding.

References

1 K. Gehring, J.-L. Leroy and M. Gueron, *Nature*, 1993, **363**, 561-565.
2 M. Gueron and J.-L. Leroy, *Curr. Opin. Struct. Biol.*, 2000, **10**, 326-331.
3 A. T. Phan and J.-L. Mergny, *Nucleic Acids Res.*, 2002, **30**, 4618-4625.
4 Y. Zhao, Y.-X. Zeng, Z.-Y. Kan, Y.-H. Hao and Z. Tan, *ChemBioChem*, 2005, **6**, 1957-1960.
5 H. Liu, and D. Liu, *Chem. Comm.*, 2009, **19**, 2625-2636.
6 C. Chen, M. Li, Y. Xing, Y. Li, C.-C. Joedecke, J. Jin, Z. Yang and D. Liu, *Langmuir*, 2012, **28**, 17743-17748.
7 T. Liedl, T. L. Sobey and C. Simmel, *Nano Today*, 2007, **2**, 36-41.
8 D. Liu and S. Balasubramian, *Angew. Chem.*, 2003, **42**, 5734-5736.
9 Y. Mao, D. Liu, S. Wang, S. Luo, W. Wang, Y. Yang, Q. Ouyang and L. Jiang, *Nucleic Acids Res.*, 2007, **35**, e33-.
10 S. Modi, M. G. Swetha, D. Goswami, G. D. Gupta, S. Mayor and Y. Krishnan, *Nat. Nanotechnol.*, 2009, **4**, 325-330.
11 F. Pu, C. Y. Wang, D. Hu, Z. Z. Huang, J. S. Ren, S. Wang, and X. G. Qu, *Mol. Biosys.*, 2010, **6**, 1928-1932.
12 Y. H. Peng, X. H. Wang, Y. Xiao, L. Y. Feng, C. Zhao, J. S. Ren and X. G. Qu, *J. Am. Chem. Soc.*, 2009, **131**, 13813-13818.
13 J. Sponer, J. Leszczynski, V. Vetterl and P. Hobza, *J. Biomol. Struct. Dyn.*, 1996, **13**, 695-706
14 N. Spackova, I. Berger, M. Egli and J. Sponer, *J. Am. Chem. Soc.*, 1998, **120**, 6147-6151.
15 J. Gallego, E. B. Golden, D. E. Stanley, and B. R. Reid, *J. Mol. Biol.*, 1999, **285**, 1039-1052.
16 T. E. Malliavin, K. Snoussi, and J.-L. Leroy, *Magn. Reson. Chem.*, 2003, **41**, 18-25.
17 T. E. Malliavin, J. Gau, K. Snoussi and J.-L. Leroy, *Biophys. J.*, 2003, **84**, 3838-3847.
18 J. Sponer and N. Spackova, *Methods*, 2007, **43**, 278-290.
19 J. Sponer, X. Cang and T. E. Cheatham III, *Methods*, 2012, **57**, 25-39.
20 S. Ahmed and E. Henderson, *Nucleic Acids. Res.*, 1992, **20**, 507-511.
21 S. Pataskar, D. Dash and S. K. Brahmachari, *J. Biomol. Struc. Dyn.*, 2001, **19**, 293-306.
22 J. Choi, S. Kim, T. Tachikawa, M. Fujitsuka and T. Majima, *J. Am. Chem. Soc.*, 2011, **133**, 16146-16153.
23 Smiatek, J.; Chen, C.; Liu, D.; Heuer, A. *J. Phys. Chem. B*, 2011, **115**, 13788-13795.
24 Smiatek, J.; Liu, D.; Heuer, A. *Curr. Phys. Chem*, 2012, **2**, 115-123.
25 Lieblein, A. L.; Krämer, M.; Dreuw, A.; Fürtig, A.; Schwalbe, H. *Angew. Chem. Int. Edt.*, 2012, **51**, 4067-4070.
26 H. A. Day, C. Huguin and Z. A. E. Waller, *Chem. Comm.*, 2013, **49**, 7696-7698.
27 J. Smiatek, D. Janssen-Mueller, R. Friederich and A. Heuer, *Physica A*, 2013, **394**, 136-144.
28 A. T. Phan, M. Gueron and J.-L. Leroy, *J. Mol. Biol.*, 2000, **299**, 123-144
29 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 12562-12566.
30 A. Laio and F. L. Gervasio, *Rep. Prog. Phys.*, 2008, **71**, 126601-126623.
31 J. Smiatek and A. Heuer, *J. Comp. Chem.*, 2011, **32**, 2084-2096.
32 B. Ensing, M. de Vivo, Z. Liu, P. Moore and M. L. Klein, *Acc. Chem. Res.*, 2005, **39**, 73-81.
33 B. Ensing, A. Laio, M. Parrinello and M. L. Klein, *J. Phys. Chem. B*, 2006, **109**, 6676-6687.
34 G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, 2002, **417**, 876-880.
35 J. Sponer, J. Leszczynski and P. Hobza, *Biopolymers*, 2002, **61**, 3-31.
36 P. Kührova, M. Otyepka, J. Sponer and P. Banas, *J. Chem. Theory Comput.*, 2014, **10**, 401-411.
37 J. Sponer, A. Mladek, N. Spackova, X. Cang, T. E. Cheatham III and S. Grimme, *J. Am. Chem. Soc.*, 2013, **135**, 97859796.