topology-dependent one-stranded intercalation mode" Angew. Chem. 2016, 128, 7567-7570, which has been published in final form at 10.1002/ange.201602145. This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.

An Organometallic Compound which Exhibits a DNA Topology-Dependent One-Stranded Intercalation Mode

Zhujun Ma[+], Giulia Palermo[+], Zenita Adhireksan[+], Benjamin S. Murray, Thibaud von Erlach, Paul J. Dyson, Ursula Rothlisberger,[*] and Curt A. Davey[*]

[*]

Dr. Z. Ma,[+] Dr. Z. Adhireksan,[+] Prof. C. A. Davey School of Biological Sciences

Nanyang Technological University 60 Nanyang Drive, Singapore 637551 (Singapore)

E-mail: davey@ntu.edu.sg

Dr. G. Palermo,[+] T. von Erlach, Prof. P. J. Dyson, Prof. U. Rothlisberger

Institut des Sciences et Ingénierie Chimiques Ecole Polytechnique Fédérale de Lausanne (EPFL) 1015 Lausanne (Switzerland)

E-mail: ursula.roethlisberger@epfl.ch

Dr. B. S. Murray

Department of Chemistry, University of Hull Cottingham Road, Hull, HU6 7RX (UK)

[+] These authors contributed equally to this work.

Abstract: Understanding how small molecules interact with DNA is essential since it underlies a multitude of pathological conditions and therapeutic interventions. Many different intercalator compounds have been studied because of their activity as mutagens or drugs, but little is known regarding their interaction with nucleosomes, the protein-packaged form of DNA in cells. Here, using crystallographic methods and molecular dynamics simulations, we discovered that adducts formed by [(η^6-THA)Ru(ethylenediamine)Cl][PF_6] (THA = 5,8,9,10-tetrahydroanthracene; RAED-THA-Cl[PF_6]) in the nucleosome comprise a novel one-stranded intercalation and DNA distortion mode. Conversely, the THA group in fact remains solvent exposed and does not disrupt base stacking in RAED-THA adducts on B-form DNA. This newly observed DNA binding mode and topology dependence may actually be prevalent and should be considered when studying covalently binding intercalating compounds.

Intercalator compounds have been extensively studied for decades because of their cytotoxic, mutagenic and therapeutic potential.[1] Intercalators typically consist of a planar aromatic group that inserts between base pairs by unstacking and unwinding the double helix. As such, these agents can disrupt DNA transactions, such as replication and other polymerase activities. In addition to purely organic intercalators, a variety of different organometallic species have been investigated, including ones in which the metal group is available for covalent bonding to nucleobases.[2] This can impose additional constraints on intercalation mode, which may also be dependent on the topological state of DNA. In eukaryotes, DNA is packaged by histone proteins into chromatin, of which the basic repeating units are nucleosomes. However, very little is known about the site selectivity and structures associated with nucleosome-intercalator interactions.[3]

In order to better understand the physiological properties of an organometallic intercalator, we conducted a combined X-ray crystallographic and computational study of the ruthenium arene ethylenediamine (RAED) compound, RAED-THA-Cl[PF_6] (Figure 1a), which reveals a surprising dimorphism in the DNA adduct structures formed by this compound. RAED-THA-Cl[PF_6] was originally

discovered as an anticancer agent that displays unusually high cytotoxicity amongst the reported monofunctional ruthenium compounds.[4] This compound is at least as cytotoxic as cisplatin, in spite of lacking its capacity to form DNA cross-links, which has been attributed to intercalation of the THA group in DNA adducts.[5] However, THA is atypical of most intercalating groups, which are constrained to remain planar due to aromatic conjugation throughout the ring system, and our study reveals the striking consequences of this distinction.

To assess the site selectivity behavior of RAED-THA-Cl[PF_6] and the nature of adduct structures formed, we first carried out binding studies with crystals of the nucleosome core particle (NCP; composed of a 145 bp DNA and the ubiquitous histone octamer).[6] We screened many different crystals for X-ray diffraction quality subsequent to a variety of RAED-THA-Cl[PF_6] treatment concentrations and incubation durations to find an optimum that yields both high binding site occupancy and high resolution. While we see the same site selectivity over different treatment strengths and durations, we conducted structural refinement with a high site occupancy dataset having reflections to 2.6 Å resolution (Tables S1, S2).

RAED-THA-$Cl^{[+]}$ is seen to associate by substitution of the chloride anion (presumably via an aqua intermediate), yielding mono-coordinated RAED-THA adducts (Figures 1, 2). There are three principal binding sites for RAED-THA, comprising two major groove adducts at symmetry-related locations of SHL ± 1.5 (SuperHelix Location, 1.5 double helical turns from the nucleosome center) and one within the minor groove at SHL + 1 (Figures 1, 2, S1–S3). Since adduct formation at the SHL + 1 site is dependent on a specific histone-DNA internucleosomal interaction of unknown physiological relevance, it may correspond to a crystal packing artifact and we have not discussed it further here. The major groove adducts are very similar between the two symmetry-related sites and entail monofunctional coordination of the ruthenium ion to the N7 nitrogen of the 5' guanine (bp ± 15) within an AGG sequence element. The THA groups in the SHL ± 1.5 adducts are situated 5' to the coordinated guanine base, whereby the distal ring stacks between the bp ± 15G and ± 16A purines and is situated nearly coplanar with the guanine base. The massive kink distortion into the minor groove at this AG = CT step (*roll* = -33°, average between SHL ± 1.5; Figure S4) prevents coplanar stacking onto both purines, thus only the THA periphery makes close van der Waals contacts with the 5' adenine base (Figures 1b, 2, S2). This conservative, one-stranded intercalation mode is fostered by the hinge at the THA center, whereby an approximate 30° bend allows interbase stacking while maintaining pseudo-octahedral coordination geometry about the Ru center. The ethylenediamine group, conversely, is situated on the 3' side of the coordinated guanine base, where it engages in H-bonding with the O6 atom[7] (Figures 2, S2).

RAED-THA adducts form on the nucleosomal DNA at a pair of guanine sites having the highest steric accessibility, which is a consequence of severe double helix deformation induced by histone binding at SHL ± 1.5.[6a,8] Nonetheless, RAED-THA reaction at SHL ± 1.5 surprisingly results in intercalation within the ±15/16 AG = CT step, as opposed to "facile" intercalation within the pre-existing (or dynamically accessible) unstacked ± 14/15 GG = CC step (Figures S2e, S5a). This yields a hitherto unseen DNA deformation in the nucleosome, with the extreme base pair unstacking and kink into the minor groove associated with the ± 15/16 step (Figures 2, S2). The origin of this peculiar DNA site selectivity and conformational remodelling is a result of steric factors (Figure S5).

In order to understand the basis for the DNA site selectivity and adduct structures displayed by the RAED-THA$^{[2+]}$ dication, we conducted molecular dynamics (MD) simulations at both the force field (FF) and hybrid molecular mechanical/quantum mechanical (QM/MM) levels.[9] We started by carrying out MD based on the crystal structure of RAED-THA-NCP, containing the two SHL ± 1.5 RAED-THA-DNA adducts (Movie S1). Analysis of the roughly 500 ns FF and 30 ps QM/MM trajectories at 300 K shows that both levels of theory yield qualitatively very similar results, which reveal that RAED-THA maintains stable, one-stranded intercalation throughout the MD simulations (Figures 3, S6). Moreover, the nature of THA stacking against the coordinated guanine base and overall geometry of the adducts between the MD and X-ray structures are very similar, which confirms the accuracy of the computational methodology (Table S3). MD simulations at 300 K further reveal that the conformational freedom of both the RAED-THA and the DNA at the adduct sites throughout the trajectories are limited, indicating that the histone-bound configuration is highly constrained (Figure S7a, Table S4).

Given the nature of conformational constraints and the degree of deformation of the double helix imposed by histone binding, we next probed whether the RAED-THA adduct structure would change if the histone-associated constraints were eliminated. For this, we isolated a 14 bp section of nucleosomal DNA (*nuc*DNA) from the RAED-THA-NCP model encompassing the adduct and conducted 550 ns FF MD (Figure 4, Movie S2). The trajectory shows that the double helix rapidly (within the first 50–60 ns) relaxes to an undistorted B-form conformation (Figure S7b), in which the base pairs are stacked and the THA group is extruded (Figure 4). Concomitantly, the THA group rotates away from the bases and remains unintercalated for the rest of the trajectory, moving relatively freely in the solvent exposed major groove. This indicates that histone-induced deformation of the double helix is required for maintaining the mode of THA intercalation observed in the RAED-THA-NCP model. The same results are obtained in the case of a double-stranded DNA (*ds*DNA) with an identical sequence (Figure 4b, Movie S3). The inability of RAED-THA to intercalate into *ds*DNA is apparently not a consequence of the lack of aromatic character for the peripheral THA ring, since we find that RAED-DHA (DHA = 9,10-dihydroanthracene)[10] adducts behave in an effectively identical fashion (Figure S8).

The mode of RAED-THA interaction with the Ru-coordinated guanine base observed here in the NCP structure is comparable to that in the small molecule crystal structure of an adduct with an isolated guanine base (9-ethylguanine).[10] On the other hand, from NMR experiments on RAED-THA adducts with a 6 bp *ds*DNA, the structure was not explicitly resolved.[5b] The absence of THA intercalation in naked B-form DNA found in our FF and QM/MM MD simulations shows that there is insufficient driving force from hydrophobic and stacking interactions (we estimate the stacking energy to be only about 10 kcal mol^{-1}; details in the Supporting Information) to compensate for the energetic penalty associated with base pair unstacking and the steric clashes caused by the bulky ligand. This is consistent with the fact that RAED-THA intercalation in the nucleosome core requires histone binding-induced unstacking of bases and DNA bending, and even as such the THA group only intercalates in a restricted, one-stranded fashion. This peculiar one- stranded intercalation mode is accompanied by extreme unstacking of the ± 15/16 AG = CT base pairs (Figure S9), whereby the associated massive kink into the minor groove is synonymous with an asymmetric separation of the base pairs that is largely localized on the major groove face. This in fact creates a wedge-shaped crevice, which is ideally suited for the RAED-THA intercalation mode we observe (Figure 2).

Because of the distinct coordination geometry of the Ru^{II} center and the presence of the THA ligand, RAED-THA does not assume the typical intercalative binding mode of nucleosomal DNA alkylating agents, which preferentially target the pre-unstacked ± 14/15 GG = CC in the native NCP structure.[3] Indeed, unlike compounds having the larger, planar group, naphthalimide, which

intercalates fully at the site of the pre-unstacked ± 14/15 GG = CC,[3] RAED-THA binding remodels the double helix deformation, with THA intercalating instead within the adjacent ± 15/16 AG = CT dinucleotide. Because of the low stacking energy for THA, a coplanar stacking within the ± 14/15 GG = CC is prevented, while a one-stranded intercalation within the adjacent ±15/16 AG = CT step is favored by a reduced double helix deformation penalty associated with intercalation at this site. Indeed, the extent of AG = CT unstacking (average rise = 6.2 Å) seen in the RAED-THA adducts is substantially less than what has been observed for naphthalimide intercalation-induced unstacking (average rise = 7.8 Å) at the adjacent GG = CC site (Figure S4). Considering moreover that none of the many RAED-THA NCP crystal incubation trials conducted showed any evidence for intercalative binding without Ru-N7 (± 15G) covalent bonding (Table S2, Figures S1, S3), which contrasts with the naphthalimide-based compounds we have studied that yield mixtures of covalently and non-covalently bound fully intercalated states,[3] metal bond formation is apparently required for stable intercalation of the THA group in the nucleosome core.

Overall, this work revealed that RAED-THA$^{[2+]}$ has an ambivalent intercalating nature by forming fundamentally different adduct structures on B-form double helix relative to other topological states of DNA with unstacked base pair deformations. The bimodal nature of RAED-THA$^{[2+]}$ activity stems from a combination of covalent binding-induced constraints by the metal center and the hinge-deformation of the THA group. This results in a bulky solvent exposed adduct on naked DNA, but yields a unique distorted double helix configuration, yet relatively solvent-recessed adduct, on the nucleosome. By potentially thwarting DNA repair processes in the cell, this may contribute to the unusually high cytotoxicity of RAED-THA-Cl[PF_6], and in fact RAED-THA adducts have been shown to be repaired with poor efficiency in vitro relative to those of either [(η^6-*p*-cymene)Ru(ethylenediamine)Cl]PF_6 or cisplatin.[5a] Future investigations of covalently binding intercalators should consider the potential for one-stranded intercalation and DNA topology-dependent binding, in particular with respect to the ubiquitous nucleosomal state of DNA in the cell.

Acknowledgements

We thank Dr. M. Wang, Dr. V. Olieric and staff at the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland). We thank Prof. L. Emsley (EPFL, Switzerland) for helpful discussions. We are grateful for funding by the Singapore Ministry of Health National Medical Research Council (grant number NMRC/1312/2011) and Ministry of Education Aca-demic Research Fund Tier 3 Programme (grant number MOE2012-T3-1-001) and the Swiss National Science Foun-dation and NCCR Chemical Biology.

[1] A. Rescifina, C. Zagni, M. G. Varrica, V. Pistar, A. Corsaro, *Eur. J. Med. Chem.* **2014**, *74*, 95 – 115.

[2] a) H. Baruah, C. G. Barry, U. Bierbach, *Curr. Top. Med. Chem.* **2004**, *4*, 1537 – 1549; b) B. M. Zeglis, V. C. Pierre, J. K. Barton, *Chem. Commun.* **2007**, *28*, 4565 – 4579.

[3] a) G. E. Davey, B. Wu, Y. Dong, U. Surana, C. A. Davey, *Nucleic Acids Res.* **2010**, *38*, 2081 – 2088; b) E. Y. Chua, G. E. Davey, C. F. Chin, P. Droge, W. H. Ang, C. A. Davey, *Nucleic Acids Res.* **2015**, *43*, 5284 – 5296.

[4] R. E. Aird, J. Cummings, A. A. Ritchie, M. Muir, R. E. Morris, H. Chen, P. J. Sadler, D. I. Jodrell, *Br. J. Cancer* **2002**, *86*, 1652 – 1657.

[5] a) O. Novakova, J. Kasparkova, V. Bursova, C. Hofr, M. Vojtiskova, H. Chen, P. J. Sadler, V. Brabec, *Chem. Biol.* **2005**, *12*, 121 – 129; b) H. K. Liu, J. A. Parkinson, J. Bella, F. Y. Wang, P. J. Sadler, *Chem. Sci.* **2010**, *1*, 258 – 270.

[6] a) M. S. Ong, T. J. Richmond, C. A. Davey, *J. Mol. Biol.* **2007**, *368*, 1067 – 1074; b) B. Wu, M. S. Ong, M. Groessl, Z. Adhir-eksan, C. G. Hartinger, P. J. Dyson, C. A. Davey, *Chem. Eur. J.* **2011**, *17*, 3562 – 3566; c) Z. Adhireksan, G. E. Davey, P. Campo-manes, M. Groessl, C. M. Clavel, H. Yu, A. A. Nazarov, C. H. Yeo, W. H. Ang, P. Droge, U. Rothlisberger, P. J. Dyson, C. A. Davey, *Nat. Commun.* **2014**, *5*, 3462.

[7] B. Spingler, D. A. Whittington, S. J. Lippard, *Inorg. Chem.* **2001**, *40*, 5596 – 5602.

[8] a) B. Wu, K. Mohideen, D. Vasudevan, C. A. Davey, *Structure* **2010**, *18*, 528 – 536; b) B. Wu, G. E. Davey, A. A. Nazarov, P. J. Dyson, C. A. Davey, *Nucleic Acids Res*. **2011**, *39*, 8200 – 8212; c) E. Y. Chua, D. Vasudevan, G. E. Davey, B. Wu, C. A. Davey, *Nucleic Acids Res*. **2012**, *40*, 6338 – 6352.

[9] E. Brunk, U. Rothlisberger, *Chem. Rev*. **2015**, *115*, 6217 – 6263.

[10] H. M. Chen, J. A. Parkinson, S. Parsons, R. A. Coxall, R. O. Gould, P. J. Sadler, *J. Am. Chem. Soc.* **2002**, *124*, 3064 – 3082.

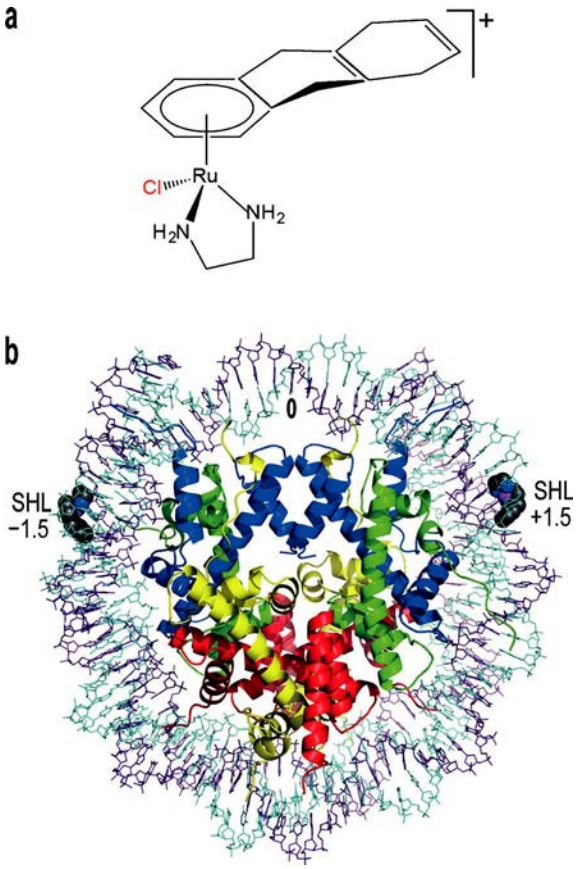

Figure 1. RAED-THA binding in the nucleosome core. a) Chemical structure of RAED-THA. b) X-ray crystal structure of RAED-THA-NCP, viewed looking down the DNA superhelical axis, with the NCP pseudo-twofold axis running vertically (through the central base pair, 0). Histone proteins are shown in blue (H3), green (H4), yellow (H2A) and red (H2B), and the two 145-nucleotide DNA strands are cyan and purple. The two symmetry-related RAED-THA adducts on the DNA are shown with space-filling representation.

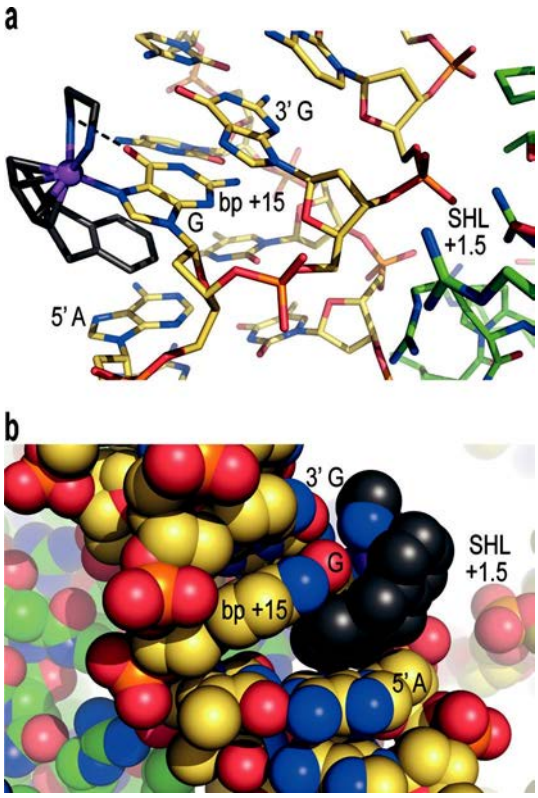

Figure 2. Detailed view of RAED-THA-DNA adduct in the nucleosome core. DNA (gold), histone proteins (green) and RAED-THA (gray) are shown in a) stick and b) space filling representation. The massive histone binding-induced kink distortion at the bp +14/15 GG =CC step in the native NCP structure (Figures S2e, S5a) is transformed into a novel extreme deformation at the bp +15/16 AG =CT step here in the adduct structure.

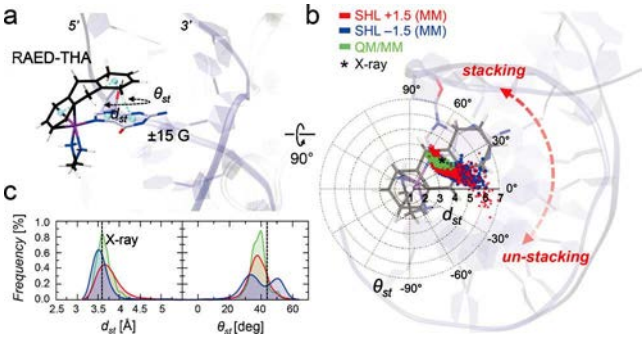

Figure 3. MD simulations of RAED-THA-NCP. RAED-THA and the coordinated ± 15G are shown as sticks, with the surrounding histones and DNA as ribbons. a) The THA stacking against the coordinated ± 15G is measured by using: 1) the distance d_{st} between the center of masses (COM) of the THA 5,8-dihydrobenzene and the ± 15G six-membered ring and 2) the dihedral angle θ_{st} between the COM of the THA arene, the THA 5,8-dihydrobenzene, the ± 15G six-membered ring and the ± 15G five-membered ring. b) 90° rotated view superposed with the polar plot of d_{st} (polar coordinate) with respect to θ_{st} (angular coordinate), from FF (red and blue for two SHL ± 1.5 sites) and QM/ MM (green) MD. X-ray values are indicated with an asterisk. A red arrow indicates the region of stacked/unstacked THA- ± 15G configurations. c) Corresponding probability distributions of d_{st} and θ_{st} (expressed in Ångström and degrees, respectively) and their respective X-ray values (dashed lines).

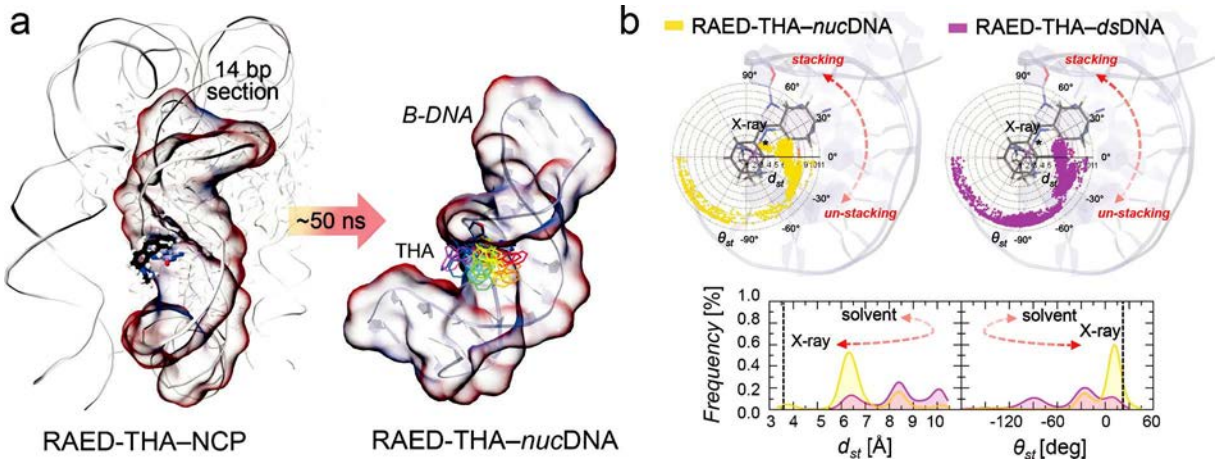

Figure 4. Results of MD simulations of a RAED-THA adduct bound to a 14 bp section of nucleosomal DNA (*nuc*DNA) and canonical double-stranded DNA (*ds*DNA; details are given in the Supporting Information). a) X-ray structure of RAED-THA-NCP showing the 14 bp section together with RAED-THA covalently bound at + 15G (left). Structure of the histone-removed DNA fragment after 50–60 ns of classical MD with an extruded RAED-THA (right; shown with different superposed configurations color-coded, red-to-violet). b) Polar plots (top) and probability distributions (bottom) of d_{st} and θ_{st} (in Ångström and degrees, respectively; defined in Figure 3), as calculated from 550 ns of classical MD of the RAED-THA adduct with a 14 bp *nuc*DNA (RAED-THA-*nuc*DNA, orange) and with canonical *ds* B-DNA (RAED-THA-*ds*DNA, magenta).