

# AT-hook peptides bind the major and minor groove of AT-rich DNA duplexes

Alyssa Garabedian<sup>1</sup>, Kevin Jeanne Dit Fouque<sup>1,3</sup>, Prem P. Chapagain<sup>2,3</sup>, Fenfei Leng<sup>1,3</sup> and Francisco Fernandez-Lima<sup>1,3,\*</sup>

<sup>1</sup>Department of Chemistry and Biochemistry, Florida International University, Miami, 33199, USA, <sup>2</sup>Department of Physics, Florida International University, Miami, 33199, USA and <sup>3</sup>Biomolecular Sciences Institute, Florida International University, Miami, 33199, USA

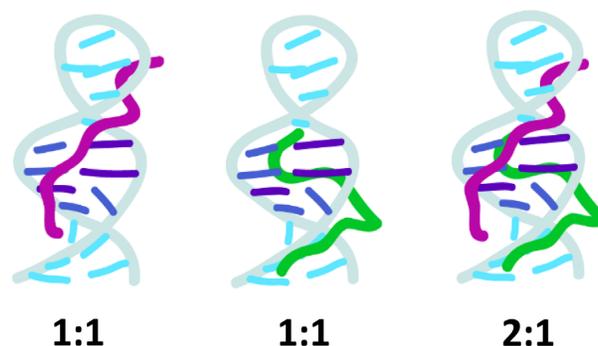
Received August 07, 2021; Revised December 30, 2021; Editorial Decision January 31, 2022; Accepted February 08, 2022

## ABSTRACT

The mammalian high mobility group protein AT-hook 2 (HMGA2) houses three motifs that preferentially bind short stretches of AT-rich DNA regions. These DNA binding motifs, known as ‘AT-hooks’, are traditionally characterized as being unstructured. Upon binding to AT-rich DNA, they form ordered assemblies. It is this disordered-to-ordered transition that has implicated HMGA2 as a protein actively involved in many biological processes, with abnormal HMGA2 expression linked to a variety of health problems including diabetes, obesity, and oncogenesis. In the current work, the solution binding dynamics of the three ‘AT-hook’ peptides (ATHPs) with AT-rich DNA hairpin substrates were studied using DNA UV melting studies, fluorescence spectroscopy, native ion mobility spectrometry-mass spectrometry (IMS-MS), solution isothermal titration calorimetry (ITC) and molecular modeling. Results showed that the ATHPs bind to the DNA to form a single, 1:1 and 2:1, ‘key-locked’ conformational ensemble. The molecular models showed that 1:1 and 2:1 complex formation is driven by the capacity of the ATHPs to bind to the minor and major grooves of the AT-rich DNA oligomers. Complementary solution ITC results confirmed that the 2:1 stoichiometry of ATHP: DNA is originated under native conditions in solution.

## GRAPHICAL ABSTRACT

### ATHP : DNA



## INTRODUCTION

The mammalian high mobility group protein AT-hook 2 (HMGA2) is a small non-histone chromosomal protein and belongs to the HMGA protein family (1,2). It has three ‘AT-hook’ DNA binding motifs and specifically binds to AT-rich DNA sequences (3–6). The ‘AT-hook’ DNA-binding motif is an 8–9 amino acid peptide with a conserved tripeptide -Arg-Gly-Arg- core surrounded by multiple positively charged amino acids lysines and/or arginines (3). HMGA2 is a multifunctional nuclear protein and is likely a key player for several important biological processes that yield certain unique phenotypes. For instance, HMGA2 is linked to obesity (7,8), diabetes (9), stem cell youth (10), and oncogenesis (11–14). HMGA2 is also associated with human height (15–17) and intelligence (18). Due to the importance of HMGA2 in adipogenesis (19) and tumorigenesis (20), HMGA2 is considered a potential therapeutic target for anticancer and anti-obesity treatments (8,21–23).

HMGA2 is an intrinsically disordered protein (IDP) (24,25) and does not have a secondary and tertiary structure in the absence of DNA. When the ‘AT-hook’ DNA binding motifs bind to AT-rich DNA sequences, they adopt a defined

\*To whom correspondence should be addressed. Tel: +1 305 348 2037; Email: fernandf@fiu.edu

structure (24,26). Specifically, the central RGR core deeply penetrates into the minor groove of AT base pairs and forms a well-defined 'AT-hook' and DNA complex (24,26). Interestingly, recent mass spectrometry and simulation studies showed that certain structures of the 'AT-hooks' are more stable and similar to the structures observed in the NMR and crystal structural studies (2,11), suggesting that the role of AT-rich DNA is to stabilize these structures or conformations.

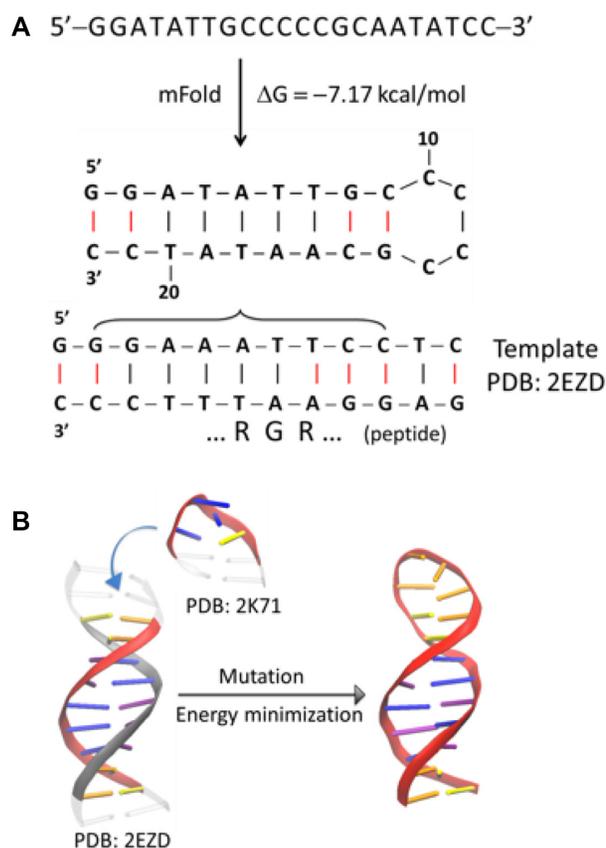
Previously, using a PCR-based systematic evolution of ligands by exponential enrichment (SELEX) experiment, we found that HMGA2 prefers binding to the following two consensus DNA sequences, 5'-ATATTGCGCAWWATT-3', and 5'-ATATTGCGCAWWATT-3', where W represents A or T (5). Intriguingly, these two HMGA2 binding sequences contain a 4-base pair GC sequence in the center. Because each 'AT-hook' binds to 5 base pairs and the minor groove of the GC sequence is crowded, one 'AT-hook' likely interacts primarily with the major groove of the center GC sequence. This hypothesis prompted us to investigate whether the 'AT-hooks' also tightly bind to the DNA major groove if the minor groove is not available. Since a variety of DNA-binding proteins carry 'AT-hook' DNA-binding motifs (27), not only will the discovery/confirmation of 'AT-hooks' binding to the DNA major groove change our view toward this DNA binding motif, but it also has great biological consequences. One implication is that HMGA proteins or DNA-binding protein carrying 'AT-hook' motifs, if minor groove is not available, can still interact with DNA through major groove and participate in nuclear activities, such as transcription, chromatin remodeling, and DNA repair.

One approach to address this hypothesis is to study how the 'AT-hook' peptides (ATHPs) derived from the three 'AT-hook' DNA binding motifs of HMGA2 bind to DNA molecules containing only one AT-rich DNA binding site. In the present work, we report on the binding dynamics of the three ATHPs of HMGA2 with AT-rich DNA hairpin substrates using DNA UV melting studies, fluorescence spectroscopy, isothermal titration calorimetry (ITC), ion mobility spectrometry – mass spectrometry (IMS-MS) and molecular modeling. Although prior studies suggested minor groove ATHP interactions, here we present, for the first time, solution- and gas-phase experimental and theoretical evidence of both ATHP major and minor DNA groove-binding capabilities.

## MATERIALS AND METHODS

### Materials and reagents

AT-rich DNA oligomers, FL876 (sequence: 5'-GGATATTGCCCCGCAATATCC-3' ( $C_{212}H_{270}N_{79}O_{130}P_{21}$ , MW 6655.1561)), FL876T1 (sequence: 5'-GGATATTGCCTCCGCAATATCC-3' ( $C_{213}H_{272}N_{78}O_{131}P_{21}$ , MW 6670.3326), and FL876T2 (sequence: 5'-GGATATTGCCTTCGCAATATCC-3' ( $C_{214}H_{273}N_{77}O_{132}P_{21}$ , MW 6685.3440)) were purchased from Eurofins Genomics (Luxembourg, Luxembourg) and used as received (Scheme 1 and Supplementary Scheme S1). These three 22 nucleotide DNA hairpins contain a 9-base pair stem including a 5 base pair AT DNA in the middle of



**Scheme 1.** Preparation of PDB template for modeling of DNA.

the stem. AT-hook peptides 1 (Lys-Arg-Gly-Arg-Gly-Arg-Pro-Arg-Lys), 2 (Pro-Lys-Arg-Pro-Arg-Gly-Arg-Pro-Lys) and 3 (Lys-Arg-Pro-Arg-Gly-Arg-Pro-Arg-Lys-Trp), which correspond to the first, second, and third 'AT-hook' motifs of HMGA2, were purchased from Advanced ChemTech Inc. (Louisville, KY) and used as received. Solvents, 1xBPE buffer (6 mM  $Na_2HPO_4$ , 2 mM  $NaH_2PO_4$ , and 1 mM  $Na_2EDTA$ , pH 7.0), and ammonium acetate salts utilized in this study were analytical grade or better and purchased from Fisher Scientific (Pittsburgh, PA). A Tuning Mix calibration standard (G24221A) was obtained from Agilent Technologies (Santa Clara, CA) and used as received.

### DNA UV melting studies

DNA UV melting curves were determined using a Cary 100 UV-Vis spectrophotometer (Agilent Technologies, Santa Clara, CA) equipped with a thermoelectric temperature-controller. DNA oligomer FL876 in 1 × BPE buffer (16 mM of  $Na^+$ ) was used for melting studies. Under this condition, the AT-rich DNA site of FL876 is in the double-stranded form at room temperature. DNA-peptide samples were prepared to a final concentration of 2  $\mu M$  by directly mixing at a molar ratio of 1:10 FL876 DNA with each ATHP peptide, followed by incubation for 60 min at room temperature to ensure equilibration. Samples were typically heated in the 20–100°C range at a rate of 1°C  $min^{-1}$ , while continuously monitoring the absorbance at 260 nm. Primary data were

transferred to the graphic program Origin (MicroCal, Inc., Northampton, MA) for plotting and analysis.

### Fluorescence measurements

Fluorescence spectra of DNA-ATHPs were acquired using a Jobin Yvin Horiba FluoroMax-3 with excitation wavelength of 355 nm. In the titration experiments, 99 nM FL876 and 50 nM Hoechst stain in  $1 \times$  BPE was titrated by increasing concentrations of ATHP 1, 2, and 3 up to 20  $\mu$ M. The fluorescence spectra were recorded from 400 to 550 nm.

### Isothermal titration calorimetry (ITC)

ITC experiments were performed using a VP-ITC titration calorimeter (MicroCal, Inc., Northampton, MA) interfaced to a personal computer. Origin 7.0 was used for data acquisition and processing. A typical ITC experiment was set up so that 10  $\mu$ L of ATHP ligand, i.e., 300  $\mu$ M ATHP1, 300  $\mu$ M ATHP2 and 150  $\mu$ M ATHP3, was injected every 120 seconds for 30 injections into a 10  $\mu$ M DNA sample (1x BPE buffer) in the sample cell.

### Native trapped ion mobility spectrometry – mass spectrometry analysis (TIMS-MS)

Details regarding the TIMS operation and specifics compared to traditional IMS can be found elsewhere (11,28–31). Briefly, mobility separation in TIMS is based on holding the ions stationary against a moving gas using an electric field. The separation in a TIMS device can be described in the center of the mass reference frame using the same principles as in a conventional IMS drift tube (32). Since mobility separation is related to the number of ion-neutral collisions (or drift time in traditional drift tube cells), the mobility separation in a TIMS device depends on the bath gas drift velocity, ion confinement and ion elution parameters. The reduced mobility,  $K$ , of an ion in a TIMS cell is described by:

$$K = \frac{V_g}{E} = \frac{A}{(V_{elution} - V_{out})}$$

where  $v_g$  and  $E$  are the velocity of the gas and the applied electric field across the TIMS analyzer region.  $V_{elution}$  is the voltage when the ions elute in the  $V_{ramp}$  sweep and  $V_{out}$  is the voltage applied at the end of the TIMS analyzer region.

A custom-built, pulled capillary nanoESI source was utilized for all the experiments. Quartz glass capillaries (O.D.: 1.0 mm and I.D.: 0.70 mm) were pulled utilizing a P-2000 micropipette laser puller (Sutter Instruments, Novato, CA) and loaded with 10  $\mu$ L aliquot of the sample solution. Sample solutions consisted of 1–10  $\mu$ M ATHP or DNA in 10 mM ammonium acetate solution at physiological pH (pH = 6.7). For the observation of the DNA-ATHP complexes, a 1:1 and 1:2 ratio of 5  $\mu$ M concentration of the DNA and AT-hook peptide (1, 2 or 3) was prepared in 10mM ammonium acetate immediately prior infusion. A typical nanoESI source voltage of +/- 600–1200 V was applied between the pulled capillary tips and the TIMS-MS instrument inlet. Ions were introduced via a stainless-steel tube (1/16  $\times$  0.020', IDEX Health Science, Oak Harbor, WA) held at room temperature into the TIMS cell.

Mobility calibration was performed using the Tuning Mix calibration standard (G24221A, Agilent Technologies, Santa Clara, CA) in positive ion mode (e.g.,  $m/z = 322$ ,  $K_0 = 1.376 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  and  $m/z = 622$ ,  $K_0 = 1.013 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ) (31). The TIMS operation was controlled using in-house software, written in National Instruments Lab VIEW, and synchronized with the maXis Impact Q-ToF acquisition program (28,29).

Reduced mobility values ( $K_0$ ) were correlated with CCS ( $\Omega$ ) using the equation:

$$\Omega = \frac{(18\pi)^{1/2}}{16} \frac{z}{(k_B T)^{1/2}} \left[ \frac{1}{m_i} + \frac{1}{m_b} \right]^{1/2} \frac{1}{K_0} \frac{1}{N^*}$$

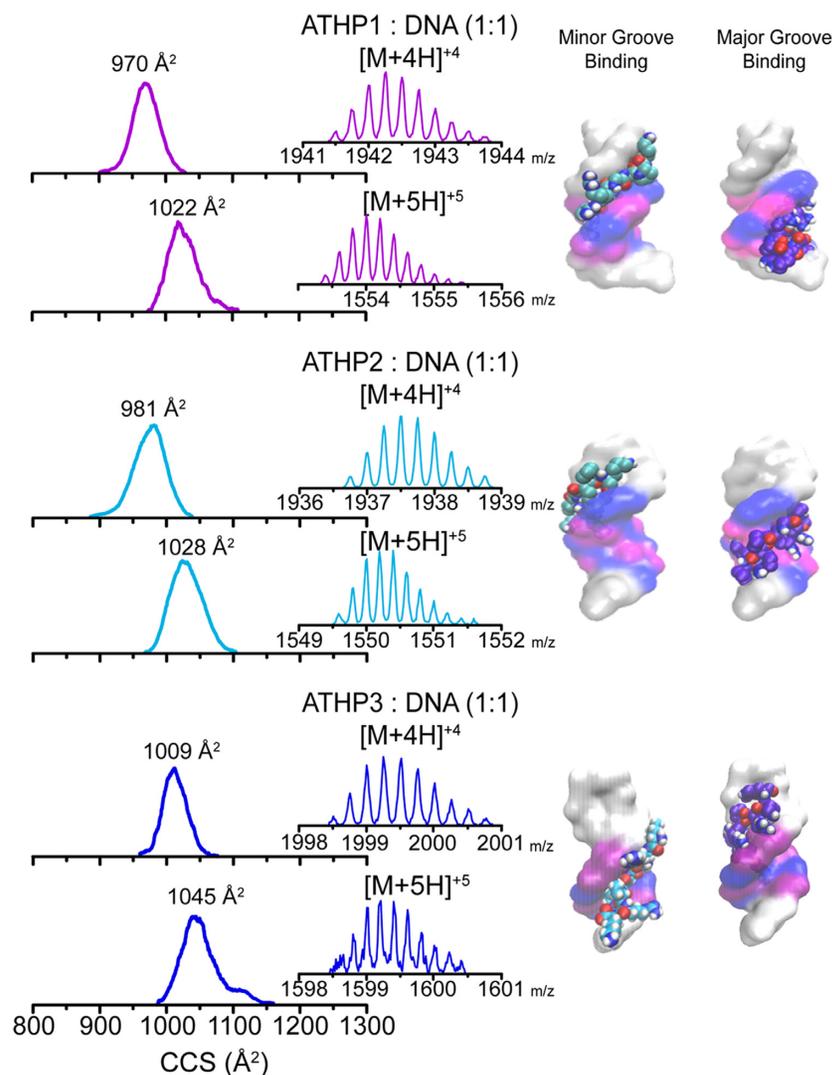
where  $z$  is the charge of the ion,  $k_B$  is the Boltzmann constant,  $N^*$  is the number density and  $m_i$  and  $m_b$  refer to the masses of the ion and bath gas, respectively (32).

### Molecular modeling

Candidate structures were proposed for the mobility bands observed in the TIMS-MS experiments. The protein data bank (PDB) entry 2EZD was used as a template for the peptide-DNA candidate structures. First, a 2-dimensional DNA hairpin structure was obtained for the DNA sequence using mFold server (33). Based on this structure, a 3-dimensional hairpin structure was then created by merging PDBs 2EZD (for the stem) with 2K71 (for the tetranucleotide loop) followed by mutations with appropriate bases to achieve the desired structure (see Scheme 1). The modeled DNA hairpin structure was optimized by energy minimization with CHARMM36 force field (34) using NAMD Molecular Dynamics package (35). Molecular docking was performed using AutoDock vina (36) to generate the DNA-ATHP complexes. Theoretical CCS were calculated using the IMoS (v1.04b) (37–39) and PSA (40) packages with nitrogen as a bath gas at ca. 300K. In the IMoS calculations, 100 total rotations were performed using the trajectory method with a Maxwell distribution. Molecular visualization was performed using Visual Molecular Dynamics software (41).

## RESULTS AND DISCUSSION

The native IMS-MS analysis of the ATHP 1, 2 or 3 in complex with the AT-rich DNA hairpin substrate (FL876 oligonucleotide that forms a stem-loop or hairpin structure and contains a single AT site for ATHP binding) produced  $[M + 4H]^{+4}$  and  $[M + 5H]^{+5}$  ions (Figure 1). The native IMS-MS analysis of equal molar sample ratios shows a complex formation with 1:1 ATHP peptide:DNA stoichiometry, accompanied by each species in their free form for all three DNA duplexes (Supplementary Figure S1-S3). Differences in the ATHPs:DNA duplexes binding were not observed depending on the hairpin sequence. Inspection of the mobility profiles of the ATHP:DNA complex shows a single mobility band for each charge state; this observation is indicative of a locked complex structure when compared to the several mobility bands observed for the ATHPs in the free form. Candidate structures based on molecular modeling were proposed based off these initial findings and sug-



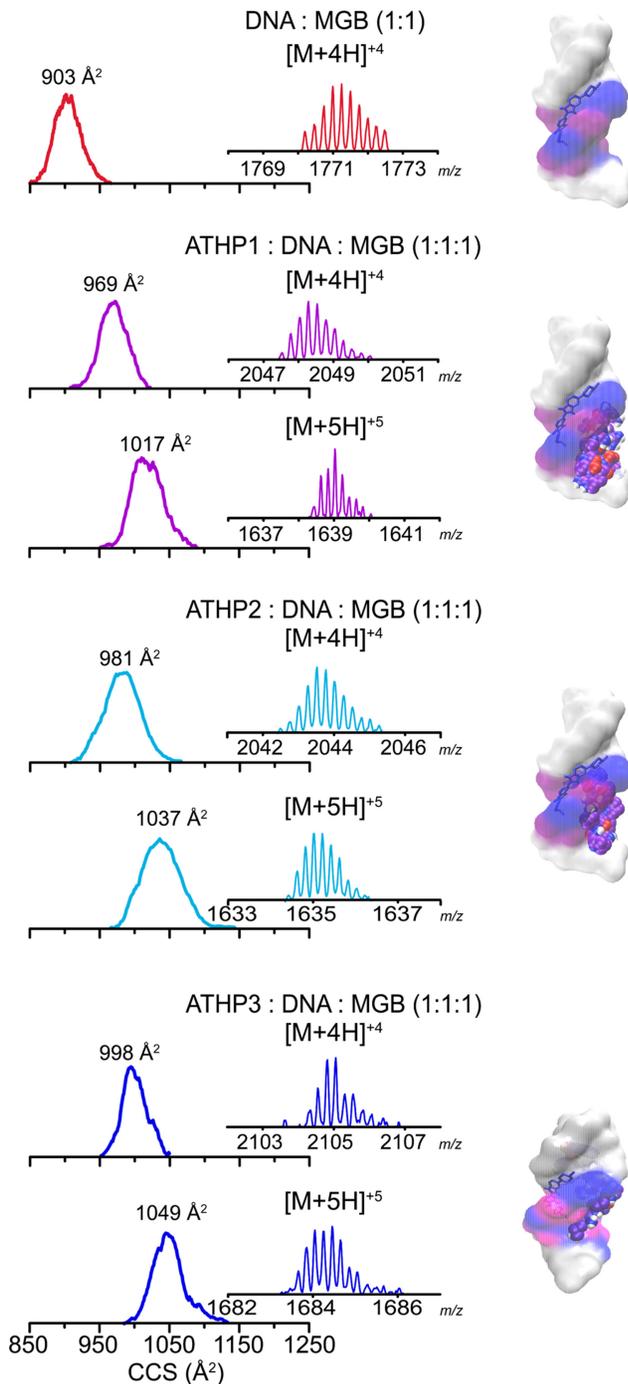
**Figure 1.** Typical mobility profiles, MS distributions (inset) and candidate structures (right) are shown for the  $[M + 4H]^{4+}$  and  $[M + 5H]^{5+}$  charge states of ATHP 1, 2, or 3 in complex with the DNA hairpin (FL876) with a ATHP:DNA 1:1 stoichiometry.

gest the likelihood of the ATHPs attaching to either the major or minor grooves of the DNA hairpin. The small difference in the CCS values of the candidate structures corresponding to the major or minor grooves ATHP:DNA 1:1 stoichiometry does not allow for their unambiguous assignment (Supplementary table S1); that is, a good agreement is observed for both candidate structures with the experimental mobility profiles. This contrasts previous assumptions in the literature (e.g., ATHPs are specific minor groove binding molecules) and subsequently motivated a series of experiments to evaluate the binding specificity of each motif.

We proceeded by using Hoechst 33528, a fluorescent dye known as a minor groove binding (MGB) compound that is highly specific to AT-rich DNA substrates and tightly binds to AT-rich DNA regions with a binding affinity in the range of  $10^8 \text{ M}^{-1}$  (42). By blocking the minor groove, ATHP attachment was restricted to the major groove of FL876 so that the presence or absence of binding could be

determined. Although displacement of the Hoechst 33258 compound in the minor groove by ATHPs is possible, we hypothesized that the majority of ATHPs would attach to the major groove, if applicable, leaving the minor groove occupied by the dye. In fact, native nESI-IMS-MS results confirmed that all three ATHPs have the ability to bind to the DNA (FL876): Hoechst complex, forming 1 : 1 : 1 molecular assemblies (Figure 2). Inspection of the mobility profiles showed a single mobility band, also indicative of a locked 1:1:1 complex structure. The molecular modeling including the MGB compound supported the hypothesis of the additional ATHP major groove binding capabilities; a good agreement was observed between the theoretical CCS values from the candidate structures of the 1:1:1 DNA(FL876):Hoechst:ATHP complexes and the experimental mobility measurements (Supplementary Table S2).

Upon confirmation of the interaction between ATHP and the major groove of FL876 DNA duplex, the peptide concentration was increased two-fold, with respect to DNA,



**Figure 2.** Typical mobility profiles, MS distributions (inset) and candidate structures (right) are shown for the  $[M + 4H]^{+1}$  and  $[M + 5H]^{+5}$  charge states of ATHP 1, 2, or 3 in complex with the DNA hairpin (FL876) and the minor groove binding (MGB) Hoechst 33258 with a ATHP: DNA : MGB 1:1:1 stoichiometry

to assess the possibility of two peptides with one DNA (bound to the major and minor grooves). An increase in the ATHP: DNA ratio to 2:1 resulted in the observation of the corresponding  $[M + 5H]^{+5}$  complex (Figure 3 and Supplementary Figures S1-S3). The small difference between the CCS values ( $<200 \text{ \AA}^2$ ) between the ATHP: DNA 1:1 and

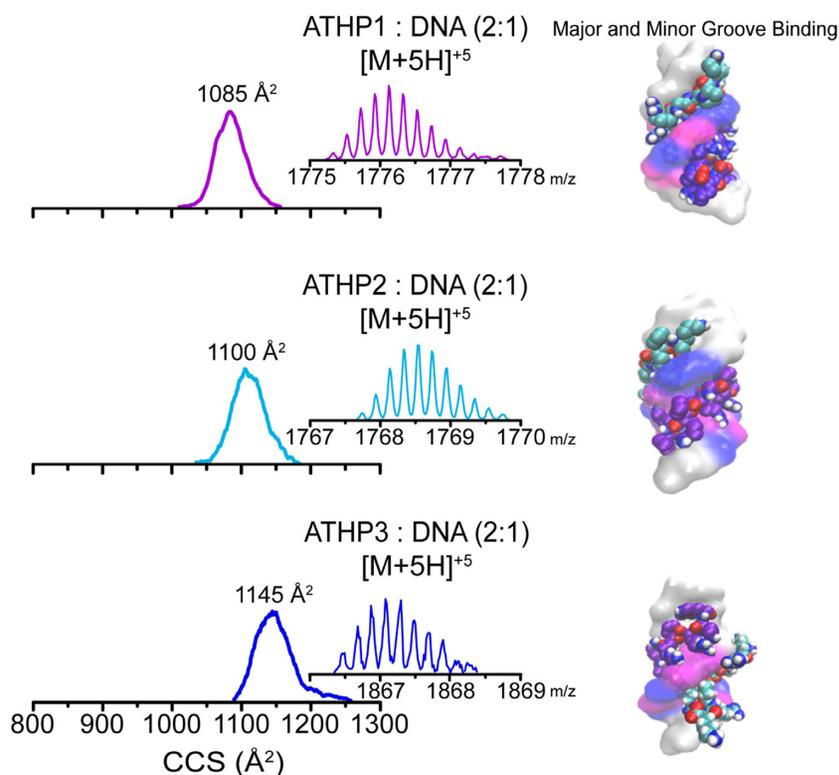
2:1 complexes suggest that the ATHP peptide(s) are mostly concealed within the DNA duplex structure. This is a common trend for the three considered DNA duplexes (Supplementary Figure S2 and S3) which suggest that there is no influence of the hairpin constituents in the ATHP-DNA binding dynamics. The proposed candidate structures provide an illustration of the ATHP:DNA (FL876) 2:1 complex; inspection of the candidate structures suggests the CCS is mostly dictated by the DNA substrate. A good agreement is also observed between the experimental CCS and those derived from the candidate structure 2:1 ATHP-DNA complexes (Supplementary Table S1). The IMS-MS results and theoretical docking studies are also in good agreement with previous structural findings using NMR and X-ray crystallography; that is, the  $-\text{Arg-Gly-Arg}-$  core of the AT-hook peptide can be buried into the minor groove of DNA, and that complexes are formed via hydrogen bonds between the oxygen atoms of thymine and/or cytosine, while Van der Waals forces govern interactions with adenines (24,26,27). The lysines and arginines outside of this core extend tightly along the DNA forming additional points of contact as a result of electrostatic interactions with the negatively charged DNA phosphate backbone (24,26,27).

Comparison of the mobility distributions of the 1:1 and 2:1 ATHP: DNA stoichiometries suggest that the ATHPs bind to the AT-rich regions of the DNA in a single 'locked' conformation and that all three ATHPs are selectively stabilized upon binding to the pre-formed FL876 hairpin (Scheme 2) (43). The schematic model, determined using the CCSs measured in the native IMS-MS, illustrates conformational changes of the peptide (i.e., unfolding) induced upon minor and/or major groove substrate binding (44). These observations suggest a disordered-to-ordered transition of HMGA2 when assemblies with DNA, where the binding occurs indistinctly via the minor or major groove.

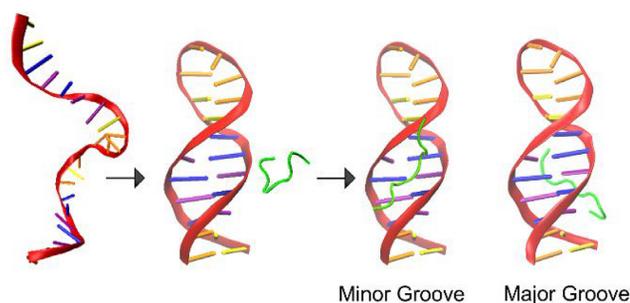
The ATHP: DNA complexes were also studied using fluorescence titration measurements, DNA melting experiments, ITC, and collision induced dissociation MS/MS (MS-CID-MS). Supplementary Figure S4 shows results of fluorescence titration experiments where increasing concentrations of ATHP1, 2, and 3 were titrated into a Hoechst 33258-FL876 complex. Since Hoechst 33258 tightly binds to the minor groove of AT-rich DNA (42,45), ATHPs should not completely displace Hoechst 33258 from the minor groove. In this way, the fluorescence intensity of the Hoechst 33258-FL876 complex should not greatly decrease. Indeed, high concentrations of ATHPs ( $20 \mu\text{M}$ ) could not fully quench the fluorescence of the Hoechst 33258-FL876 complex, indicating that ATHPs cannot completely remove Hoechst 33258 from the minor groove. This result is consistent with our previous IMS-MS studies.

Hairpin formation was confirmed by a sharp DNA UV melting transition at  $59^\circ\text{C}$  in  $1 \times \text{BPE}$  (see Figure 4A). The ATHPs 1- 3 showed a melting temperature ( $T_m$ ) increase, indicative of binding between each peptide and hairpin. The thermal stabilization of FL876 by ATHP 1 and ATHP 2 is greater than that caused by ATHP 3, indicating that ATHP1 and 2 may bind more tightly to FL876 than ATHP3 does.

The ITC results also supported the dual binding site hypothesis. Figure 4B shows the ITC results in which all can be fitted to a two-site model. Binding of ATHP1 to FL876



**Figure 3.** Typical mobility profiles, MS distributions (inset) and candidate structures (right) are shown for the  $[M + 4H]^+1$  and  $[M + 5H]^+5$  charge states of ATHP 1, 2, or 3 in complex with the DNA hairpin (FL876) with a ATHP: DNA 2:1 stoichiometry.

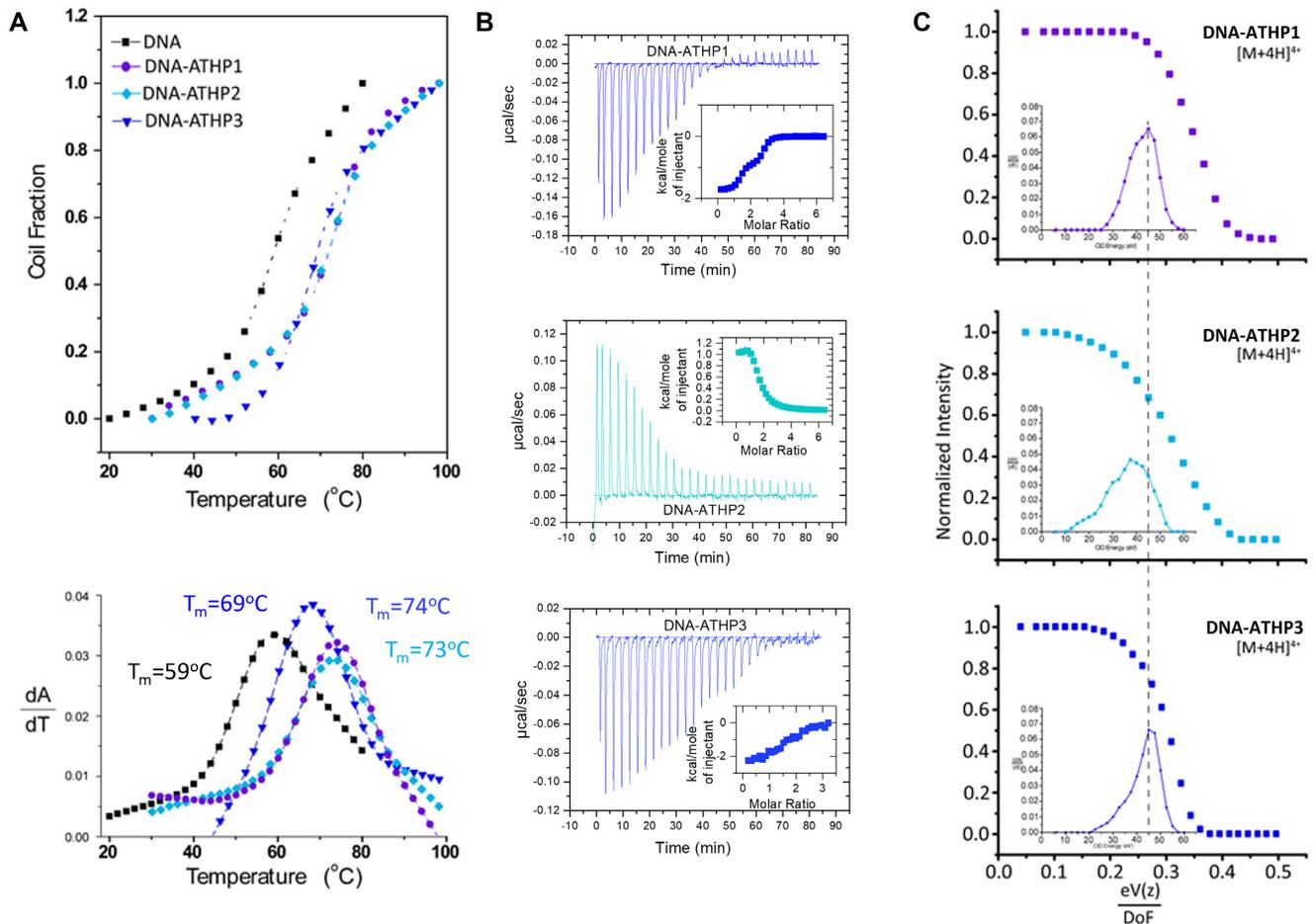


**Scheme 2.** Mechanism of ATHP attachment to DNA showing the pre-formed hairpin prior to peptide unfolding upon binding.

results in  $n1$  of  $1.3 \pm 0.1$ ,  $K1$  of  $1.9 \pm 1.6 \times 10^8 \text{ M}^{-1}$ ,  $\Delta H1$  of  $-1.7 \pm 0.7 \text{ kcal/mol}$ ,  $n2$  of  $1.4 \pm 0.1$ ,  $K2$  of  $4.0 \pm 2.4 \times 10^6 \text{ M}^{-1}$ , and  $\Delta H1$  of  $-0.9 \pm 0.1 \text{ kcal/mol}$ ; binding of ATHP2 to FL876 results in  $n1$  of  $0.8 \pm 0.5$ ,  $K1$  of  $3.9 \pm 2.4 \times 10^5 \text{ M}^{-1}$ ,  $\Delta H1$  of  $+1.6 \pm 0.7 \text{ kcal/mol}$ ,  $n2$  of  $0.7 \pm 0.3$ ,  $K2$  of  $9.4 \pm 4.1 \times 10^6 \text{ M}^{-1}$ , and  $\Delta H1$  of  $+1.0 \pm 0.1 \text{ kcal/mol}$ ; and binding of ATHP3 to FL876 results in  $n1$  of  $1.3 \pm 0.4$ ,  $K1$  of  $1.1 \pm 1.3 \times 10^7 \text{ M}^{-1}$ ,  $\Delta H1$  of  $-2.4 \pm 0.2 \text{ kcal/mol}$ ,  $n2$  of  $1.1 \pm 0.2$ ,  $K2$  of  $1.1 \pm 1.0 \times 10^6 \text{ M}^{-1}$ , and  $\Delta H1$  of  $-1.1 \pm 0.7 \text{ kcal/mol}$ . The binding enthalpy values are quite small, which gives significant uncertainty to the estimate of the binding affinities. According to our previous studies, the binding affinity of ATHPs to AT-rich DNA sequences should be in the range of  $10^6$  to  $10^7 \text{ M}^{-1}$  (46). Interestingly, binding of ATHP1 and 3 to FL876 is an exothermic reac-

tion. In contrast, binding of ATHP2 to FL876 is an endothermic reaction. These results are consistent with our previously published results (46,47).

Complementary collision induced activation experiments (MS-CID-MS) on the  $[M + 4H]^+4$  charge state of ATHP and FL876 1:1 complexes, were used to evaluate the ATHP-DNA binding affinities (Figure 4C). As the collision energy increases, a decrease in the molecular ion signal corresponding to the complex was observed; notice that collision energy has been normalized to the charge and calculated per degree of freedom to account for mass difference and charge between the experiments, a consideration supported through previous work by Clemmer's and co-workers (48). Comparison between the CID profiles of the ATHP 1–3: FL876 complexes shows that the weakest binding profile is observed for ATHP 2: FL876, followed by ATHP 3: FL876 complex, and ATHP 1: FL876 complex. We interpret these CID profiles in that as the collision energy increases, multiple collisions with the collision gas (nitrogen in this case) leads to an increase in the vibrational modes of the complex eventually leading to dissociation into the main constituents (i.e., ATHP and FL876). That is, as the CID energy increases the internal energy increases until a threshold is achieved, leading to complex dissociation. The increased energy needed to begin dissociation of ATHP 1: FL876 and ATHP 3: FL876 distribution suggests that two distinct binding sites (e.g., minor and major groove) are occupied. The more immediate dissociation of the ATHP 2-FL876 complex suggest a predominant single binding site or two weaker attachment sites. The CID profiles combined



**Figure 4.** DNA UV melting curves for DNA hairpin (FL876) and DNA hairpin (FL876) complexed with ATHPs 1, 2 or 3 (A). ITC results determining the binding stoichiometry and affinity of ATHP 1, 2 and 3 to both the major and minor groove of FL876 DNA (B) and Collision induced dissociation for FL876 complexed with ATHP1, ATHP2 or ATHP3 (charge state and degrees of freedom were considered) with the curve derivative shown as an inset (C).

with the fluorescent titration and ITC data indicate that ATHP 1, 2 and 3 can bind hairpin DNA and occupy both minor and major grooves.

## CONCLUSIONS

These studies evaluated the binding dynamics of ATHP 1, 2 or 3 of HMAG2 in complex with an AT-rich DNA hairpin duplex (FL876). Results showed that ATHP 1–3 can associate with the AT-rich DNA to form a complex in a ‘locked’ mechanism enabled by interactions of the –Arg-Gly-Arg–tripeptide core and the AT-rich regions of the DNA duplex. The use of DNA duplexes with varying hairpin constituents showed no differences in the ATHPs :DNA complex formation. The native IMS-MS, combined with the fluorescent titration and ITC data indicate that ATHP 1, 2 and 3 can bind to DNA duplexes and occupy both minor and major grooves in a 1:1 and 2:1 ATHP: DNA stoichiometry. Molecular modeling supported the observation of the ‘locked’ complex structure, where the ATHPs are likely to bind to a pre-folded DNA duplex structure in the minor and major grooves, indistinctively. Complementary ATHP 1–3: FL876 complex analysis using MS-CID-MS showed

that the weakest binding profile is observed for ATHP 2 : FL876, followed by ATHP 3 : FL876 complex, and ATHP 1: FL876 complex.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors acknowledge the financial support from the National Science Foundation Division of Chemistry, under CAREER award CHE-1654274 with co-funding from the Division of Molecular and Cellular Biosciences, from the National Institutes of General Medicine (R01GM134247) to FFL and from the National institute of Allergy and Infectious Diseases (1R21AI125973) to FL. AG would like to acknowledge the 2017–2018 Dissertation Year Fellowship from Florida International University.

*Author Contributions:* The manuscript was written through contributions of all authors.

## FUNDING

National Science Foundation (CHE-1654274), National Institutes of General Medicine (R01GM134247) to FFL, FIU 2017–2018 Dissertation Year Fellowship to AG, and National Institute of Allergy and Infectious Diseases (1R21AI125973) to FL.

*Conflict of interest statement.* None declared.

## REFERENCES

- Bustin, M. (2001) Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem. Sci.*, **26**, 152–153.
- Su, L., Deng, Z. and Leng, F. (2020) The mammalian high mobility group protein at-hook 2 (HMGA2): biochemical and biophysical properties, and its association with adipogenesis. *Int. J. Mol. Sci.*, **21**, 3710.
- Reeves, R. and Nissen, M.S. (1990) The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J. Biol. Chem.*, **265**, 8573–8582.
- Solomon, M.J., Strauss, F. and Varshavsky, A. (1986) A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. *Proc. Natl. Acad. Sci. USA*, **83**, 1276–1280.
- Cui, T. and Leng, F. (2007) Specific recognition of AT-rich DNA sequences by the mammalian high mobility group protein AT-hook 2: a SELEX study. *Biochemistry*, **46**, 13059–13066.
- Chen, B., Young, J. and Leng, F. (2010) DNA bending by the mammalian high mobility group protein at hook 2. *Biochemistry*, **49**, 1590–1595.
- Zhou, X., Benson Kf Fau - Ashar, H.R., Ashar Hr Fau - Chada, K. and Chada, K. (1995) Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature*, **376**, 771–774.
- Anand, A. and Chada, K. (2000) In vivo modulation of hmgic reduces obesity. *Nat. Genet.*, **24**, 377–380.
- Alkayyali, S., M.L., Deshmukh, H., Ahlqvist, E., Colhoun, H., Isomaa, B., Rossing, P., Groop, L. and Lyssenko, V. (2013) Common variant in the HMGA2 gene increases susceptibility to nephropathy in patients with type 2 diabetes. *Diabetologia*, **56**, 323–329.
- Copley, M.R., Babovic, S., Benz, C., Knapp, D.J.H.F., Beer, P.A., Kent, D.G., Wohrer, S., Treloar, D.Q., Day, C., Rowe, K. *et al.* (2013) The Lin28b-let-7-Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. *Nat. Cell Biol.*, **15**, 916–925.
- Schenk, E.R., Ridgeway, M.E., Park, M.A., Leng, F. and Fernandez-Lima, F. (2013) Isomerization kinetics of at hook decapeptide solution structures. *Anal. Chem.*, **86**, 1210–1214.
- Young, A.R. and Narita, M. (2007) Oncogenic HMGA2: short or small? *Genes Dev.*, **21**, 1005–1009.
- Lee, Y.S. and Dutta, A. (2007) The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev.*, **21**, 1025–1030.
- Morishita, A., MR, Z., Mito, A., Sankarasharma, D., Szabolcs, M., Okada, Y., D'Armiento, J. and Chada, K. (2013) HMGA2 is a driver of tumor metastasis. *Cancer Res.*, **73**, 4289–4299.
- Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R.B., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B. *et al.* (2007) A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat. Genet.*, **39**, 1245–1250.
- Horikoshi, M., Yaghootkar, H., Mook-Kanamori, D.O., Sovio, U., Taal, H.R., Hennig, B.J., Bradfield, J.P., St Pourcain, B., Evans, D.M., Charoen, P. *et al.* (2013) New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat. Genet.*, **45**, 76–82.
- Yang, T.L., Y, G., Zhang, L.-S., Tian, Q., Yan, H., Guo, Y.-F. and Deng, H.-W. (2010) HMGA2 is confirmed to be associated with human adult height. *Annu Hum Genet.*, **74**, 11–16.
- Stein, J.L., Medland, S.E., Vasquez, A.A., Hibar, D.P., Senstad, R.E., Winkler, A.M., Toro, R., Appel, K., Bartecek, R., Bergmann, O. *et al.* (2012) Identification of common variants associated with human hippocampal and intracranial volumes. *Nat. Genet.*, **44**, 552–561.
- Ashar, H.R., Chouinard, R.A. Jr, Dokur, M. and Chada, K. (2010) In vivo modulation of HMGA2 expression. *Biochim. Biophys. Acta*, **1799**, 55–61.
- Fusco, A. and Fedele, M. (2007) Roles of HMGA proteins in cancer. *Nat. Rev. Cancer*, **7**, 899–910.
- Alonso, N., Guillen, R., Chambers, J.W. and Leng, F. (2015) A rapid and sensitive high-throughput screening method to identify compounds targeting protein-nucleic acids interactions. *Nucleic Acids Res.*, **43**, e52.
- Su, L., Bryan, N., Battista, S., Freitas, J., Garabedian, A., D'Alessio, F., Romano, M., Falanga, F., Fusco, A., Kos, L. *et al.* (2020) Identification of HMGA2 inhibitors by alphascreen-based ultra-high-throughput screening assays. *Sci. Rep.*, **10**, 18850.
- Thuault, S., Tan, E.J., Peinado, H., Cano, A., Heldin, C.H. and Moustakas, A. (2008) HMGA2 and smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition. *J. Biol. Chem.*, **283**, 33437–33446.
- Huth, J.R., Bewley, C.A., Nissen, M.S., Evans, J.N., Reeves, R., Gronenborn, A.M. and Clore, G.M. (1997) The solution structure of an HMGI(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.*, **4**, 657–665.
- Cui, T., Joynet, S., Morillo, V., Baez, M., Hua, Z., Wang, X. and Leng, F. (2007) Large scale preparation of the mammalian high mobility group protein A2 for biophysical studies. *Protein Pept. Lett.*, **14**, 5.
- Fonfria-Subiros, E., Acosta-Reyes, F., Saperas, N., Pous, J., Subirana, J.A. and Campos, J.L. (2012) Crystal structure of a complex of DNA with one AT-hook of HMGA1. *PLoS One*, **7**, e37120.
- Aravind, L. and Landsman, D. (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.*, **26**, 4413–4421.
- Fernandez-Lima, F.A., Kaplan, D.A. and Park, M.A. (2011) Note: integration of trapped ion mobility spectrometry with mass spectrometry. *Rev. Sci. Instr.*, **82**, 126106.
- Fernandez-Lima, F.A., Kaplan, D.A., Suetering, J. and Park, M.A. (2011) Gas-phase separation using a trapped ion mobility spectrometer. *Int J Ion Mobil Spectrom.*, **14**, 93–98.
- Schenk, E.R., Mendez, V., Landrum, J.T., Ridgeway, M.E., Park, M.A. and Fernandez-Lima, F.A. (2014) Direct observation of differences of carotenoid polyene chain cis/trans isomers resulting from structural topology. *Anal. Chem.*, **86**, 2019–2024.
- Fernandez, D.R., DeBord, J.D., Ridgeway, M.E., Kaplan, D.A., Park, M.A. and Fernandez-Lima, F.A. (2014) Ion dynamics in a trapped ion mobility spectrometer. *Analyst*, **139**, 1913–1921.
- McDaniel, E.W. and Mason, E.A. (1973) In: *Mobility and diffusion of ions in gases*. John Wiley and Sons, Inc., NY.
- Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
- Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I. *et al.* (2010) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.*, **31**, 671–690.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L. and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.
- Trott, O. and Olson, A.J. (2010) AutoDock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.*, **31**, 455–461.
- Larriba, C. and Hogan, C.J. (2013) Ion mobilities in diatomic gases: measurement versus prediction with non-specular scattering models. *J. Phys. Chem. A*, **117**, 3887–3901.
- Larriba, C. and Hogan, C.J. (2013) Free molecular collision cross section calculation methods for nanoparticles and complex ions with energy accommodation. *J. Comput. Phys.*, **251**, 344–363.
- Ouyang, H., Larriba-Andaluz, C., Oberreit, D. and Hogan, C. Jr (2013) The collision cross sections of iodide salt cluster ions in air via differential mobility analysis-mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **24**, 1833–1847.
- Bleilholder, C., Wyttenbach, T. and Bowers, M.T. (2011) A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). *Method. Int. J. Mass Spectrom.*, **308**, 1–10.

41. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
42. Haq, I., Ladbury, J.E., Chowdhry, B.Z., Jenkins, T.C. and Chaires, J.B. (1997) Specific binding of hoechst 33258 to the d(CGCAAATTTGCG)<sub>2</sub> duplex: calorimetric and spectroscopic studies. *J. Mol. Biol.*, **271**, 244–257.
43. Creaser, C.S., Griffiths, J.R., Bramwell, C.J., Noreen, S., Hill, C.A. and Thomas, C.L.P. (2004) Ion mobility spectrometry: a review. Part 1. Structural analysis by mobility measurement. *Analyst*, **129**, 984–994.
44. Politis, A., Park, A.Y., Hall, Z., Ruotolo, B.T. and Robinson, C.V. (2013) Integrative modelling coupled with ion mobility mass spectrometry reveals structural features of the clamp loader in complex with single-stranded DNA binding protein. *J. Mol. Biol.*, **425**, 4790–4801.
45. Loontjens, F.G., Regenfuss, P., Zechel, A., Dumortier, L. and Clegg, R.M. (1990) Binding characteristics of hoechst 33258 with calf thymus DNA, poly[d(A-T)], and d(CCGGAATTCCGG): multiple stoichiometries and determination of tight binding with a wide spectrum of site affinities. *Biochemistry*, **29**, 9029–9039.
46. Cui, T., Wei, S., Brew, K. and Leng, F. (2005) Energetics of binding the mammalian high mobility group protein HMG2 to poly(dA-dT)<sub>2</sub> and poly(dA)-poly(dT). *J. Mol. Biol.*, **352**, 629–645.
47. Joynt, S., Morillo, V. and Leng, F. (2009) Binding the mammalian high mobility group protein AT-hook 2 to AT-rich deoxyoligonucleotides: enthalpy-entropy compensation. *Biophys. J.*, **96**, 4144–4152.
48. Pierson, N.A. and Clemmer, D.E. (2015) An IMS-IMS threshold method for semi-quantitative determination of activation barriers: interconversion of proline cis ↔ trans forms in triply protonated bradykinin. *Int. J. Mass Spectrom.*, **377**, 646–654.