2D-IR Spectroscopy Shows that Optimised DNA Minor Groove Binding of Hoechst33258 Follows an Induced Fit Model

Lennart A. I. Ramakers[†], Gordon Hithell[†], John J. May[§], Gregory M. Greetham,[‡] Paul M. Donaldson[‡], Michael Towrie,[‡], Anthony W. Parker[‡], Glenn A. Burley[§] and Neil T. Hunt*[†]*

[†] Department of Physics, University of Strathclyde, SUPA, 107 Rottenrow East, Glasgow G4 0NG, U.K.

[‡] Central Laser Facility, Research Complex at Harwell, STFC Rutherford Appleton Laboratory, Harwell Oxford, OX11 0QX, U.K.

[§] Department of Pure and Applied Chemistry, WestCHEM, University of Strathclyde, 295 Cathedral Street, Glasgow G1 1XL, U.K.

[*] E-mail neil.hunt@strath.ac.uk; Tel +44 (0)141 5483763 (N.T.H.).

Abstract

The induced fit binding model describes a conformational change occurring when a small molecule binds to its biomacromolecular target. The result is enhanced non-covalent interactions between ligand and biomolecule. Induced fit is well-established for small molecule-protein interactions, but its relevance to small molecule-DNA binding is less clear. We investigate the molecular determinants of Hoechst33258 binding to its preferred A-tract sequence relative to a sub-optimal alternating A-T sequence. Results from 2-dimensional infrared spectroscopy, which is sensitive to H-bonding and molecular structure changes, show that Hoechst33258 binding results in loss of minor groove spine of hydration in both sequences, but an additional perturbation of the base propeller twists occurs in the A-tract binding region. This induced fit maximizes favourable ligand-DNA enthalpic contributions in the optimal binding case and demonstrates that controlling the molecular details that induce subtle changes in DNA structure may hold the key to designing next-generation DNA-binding molecules.

Introduction:

Deoxyribonucleic acid (DNA) is the fundamental repository of genetic information used in the majority of organisms in nature.[1] A structural hallmark of DNA is the anti-parallel helix where the major and minor grooves provide sites for sequence-selective binding of proteins and small molecules. Molecular recognition of double-stranded DNA (dsDNA) sequences by transcription factors for example is essential for the initiation of transcription. Furthermore, DNA-binding small molecules such as minor groove binders (MGBs) can perturb various processes associated with gene expression,[2] making them excellent candidates for the design of sequence-selective probes of DNA function in cells and potentially novel therapeutics. [3-5] A comprehensive set of guiding principles for the rational design of MGBs to target DNA sequences in a highly sequence-selective manner has yet to emerge however, as a consequence of the complex combination of competing enthalpic and entropic contributions from H-bonding, van der Waals forces and changes in hydration of the DNA and ligand.

An archetypal exemplar of this problem is the *bis*-benzimidazole family of ligands, including Hoechst33258 (H33258), which show preferential binding to AT-rich dsDNA. Spectroscopic and DNA-footprinting studies reveal that H33258 binds to dsDNA in a 1:1 stoichiometry and exhibits a 22-fold binding preference for A-tract dsDNA (*e.g.,* 5'-A_3T_3) over an alternating sequence (*e.g.,* 5'-ATATAT).[6,7,8,9] The current reasoning for this sequence preference is that the A-tracts possess a narrower minor groove with a well-defined spine of hydration. The release of this water upon ligand binding is linked to the observed entropic driving force for H33258 binding.[10,11] In contrast, crystallography suggests that the hydration of the alternating sequence minor groove is not as strongly ordered.[12-14]

Overall, H33258 binding has been found to be endothermic, but favourable enthalpic contributions arise from hydrophobic interactions between the minor groove wall and the

ligand. Direct H-bonding interactions between the ligand and dsDNA are not believed to be an important stabilising factor.[15-23] This balance in favour of entropic contributions has led to widespread claims that complex formation follows a 'rigid body' model with no significant structural change to the dsDNA upon ligand binding. This however neglects changes in base orientation upon binding revealed by crystallography[15-19] and NMR[20-23] studies. Indeed, NMR has revealed that the minor groove of A-tract dsDNA narrows upon ligand binding, rather than being inherently narrow, prompting suggestion of a significant contribution from an induced fit mechanism,[24] but unfortunately a direct comparison of optimal and sub-optimal binding was not performed.

The induced fit process, in which the biomolecule changes structure to better accommodate the ligand, is often found in binding to proteins, but rarely invoked in small-molecule-DNA binding. Further study is thus required to develop a clear picture of the molecular determinants controlling the sequence selectivity of ligand binding to the minor groove of target dsDNA sequences. We believe that two-dimensional infrared (2D-IR)[25-27] has sufficient sensitivity to differentiate the structural details that are central to the binding of H33258[28] to a preferred A-tract sequence, d(GGAAATTTGC)$_2$, (A_3T_3), from those involved in sub-optimal sequence binding, of which d(GGATATATGC)$_2$, $(AT)_3$ is our chosen exemplar.

2D-IR offers the advantage of being able to probe structure, intermolecular interactions and hydration by directly measuring the coupling and dynamics of vibrational modes. [25-27] It has extended our understanding of the nature of the vibrational modes of DNA bases,[29] Watson-Crick (W-C) base pairing [30,31] and base stacking [32] as well as revealing energy transfer mechanisms between bases and backbone[33,34] and the interactions of water with DNA.[35] Most recently, temperature-jump 2D-IR reported sequence-dependent pre-melting dynamics.[36]

Here, we show that the distinct difference between H33258 binding to A-tract and alternating sequence DNA is the loss of the ordered propeller twist arrangement of the base pairs present in the uncomplexed A-tract. This results in a subtle but distinct conformational change within the minor groove to accommodate the ligand, maximising favourable enthalpic contributions to binding. The uniquely ordered A-tract sequence also governs the formation of the spine of hydration and so its removal contributes to the entropic gain from releasing hydration water. Although the results for binding to the alternating sequence are similar, the emphasis is shifted toward loss of the spine of hydration, suggesting that the induced fit effect is crucial to the observed difference in binding affinities.

Experimental

Lyophilised, salt-free DNA oligonucleotides were obtained from Eurofins; H33258, D_2O, DMSO, monobasic and dibasic sodium phosphate were obtained from Sigma-Aldrich. All chemicals were used without further purification. All samples were prepared using pD7 phosphate buffer solution to a final duplex:H33258 ratio of 1:1 and annealed at 90 °C for 10 minutes. For all IR measurements, samples were held between two CaF_2 windows separated by a polytetrafluoroethylene spacer of 50 μm thickness. FTIR measurements were carried out using a Bruker Vertex 70 spectrometer at a resolution of 1 cm^{-1} with sample concentrations of 2.5 mM (A_3T_3 duplex/H-A_3T_3 complex) or 5 mM ($(AT)_3$ duplex/H-$(AT)_3$ complex). 2D-IR spectra were collected using the ULTRA FT-2D-IR spectrometer.[37,38] The IR pulses used had a temporal duration of ~ 100fs; a center frequency of 1650 cm^{-1} and a bandwidth of ~300 cm^{-1}, at a repetition rate of 10 kHz. FT-2D-IR measurements were carried out at concentrations of 1.25 mM (A_3T_3 duplex/H-A_3T_3 complex) or 2.5 mM ($(AT)_3$ duplex/H-$(AT)_3$ complex). Further details are given in Supporting Information.

Results and Discussion:

The FTIR spectra of the A_3T_3 DNA sequence and its complex with H33258 (H-A_3T_3) are shown in Fig.1(a). The corresponding spectra of $(AT)_3$ and H-$(AT)_3$ are shown in Fig.2(a). The IR spectra of both DNA sequences show four peaks (Fig.1(a), Fig.2(a), red), these are indicated by grey dashed lines in the figures and listed in Table 1. The results are consistent with previous studies and peaks are assigned by reference to spectra of DNA duplexes containing exclusively GC or AT base pairs, shown in Figs.1(d), 2(d) and S1.[29,32,33,39] The assignments are summarized in Table 1. The peaks in the spectra of the two DNA sequences are dominated by AT modes, though contributions from GC base pairs, which are less intense, do influence the magnitude of the absorption between the peaks. This spectral congestion can be unraveled by 2D-IR methods, which report vibrational coupling patterns in the off-diagonal region of the spectrum.

The changes to the FTIR spectra upon formation of the H-A_3T_3 and H-$(AT)_3$ complexes (Fig.1(a), Fig. 2(a), black) are displayed via difference spectra (Fig.1(a), Fig. 2(a), blue). Since H33258 exhibits no vibrational modes in this spectral region (Fig.S2), these are assigned to modifications of the DNA upon binding.

The H-A_3T_3 sequence shows three main features in the difference IR spectrum (Fig.1(a)). Two of these, near 1665 cm^{-1} and 1700 cm^{-1} (blue arrows), consist of a negative peak located to the lower frequency side of a positive peak, suggesting that the AT_{4S} and AT_{2S} modes shift to higher wavenumber. In the case of the AT_{2S} peak near 1690 cm^{-1}, the presence of the new component is visible as a shoulder on the high frequency side of the peak in Fig.1(a). The third feature (red arrow) has reversed positive/negative contributions, consistent with a shift of the $A_{R1}T$ mode to lower frequency.

The features in the FTIR difference spectrum obtained for the H-$(AT)_3$ complex (Fig.2(a)) are smaller and less well-defined than those for the complex with the A-tract sequence (Fig.1(a)).

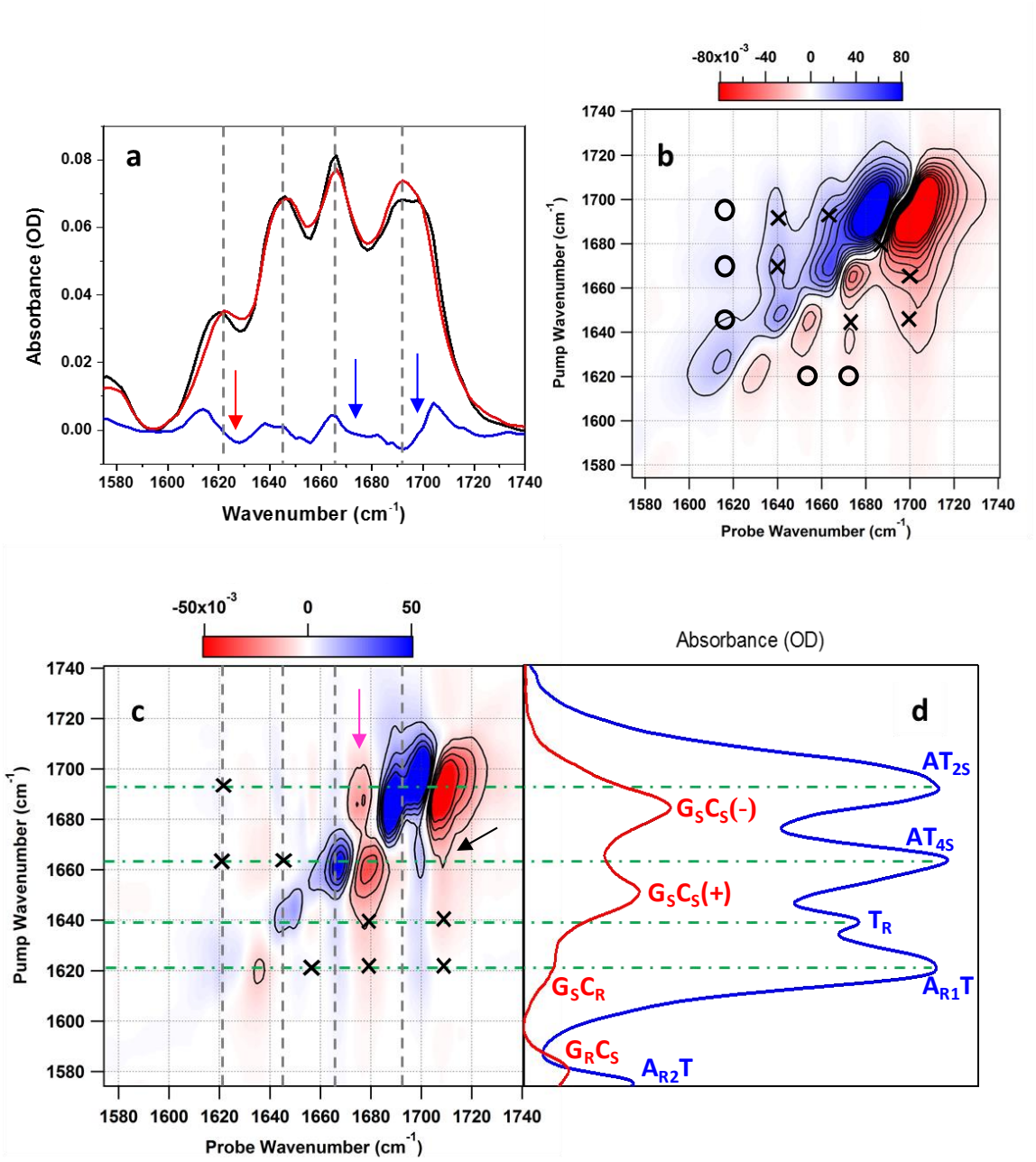

Figure 1: *a) FTIR spectra of A_3T_3 DNA with (black) and without (red) H33258. Blue spectrum shows the binding-induced difference FTIR spectrum (complex-free sequence). b) 2D-IR spectrum of H-A_3T_3 complex. Crosses mark off-diagonal peaks assigned to coupling of modes primarily located on T-base. Circles show off-diagonal peaks assigned to coupling of modes on A and T bases induced by W-C base pairing. c) Ligand binding induced 2D-IR difference spectrum. Crosses show locations of small off-diagonal features. d) FTIR spectra of GC and AT-only DNA sequences to show the relative (weighted per base) magnitudes of peaks and peak positions to aid assignment of the IR spectrum of the A_3T_3 sequence, see also Fig S1).*

Table 1: Assignment of peaks in the IR spectra of A_3T_3 and $(AT)_3$ DNA duplexes, showing comparison with AT and GC-only sequences.

Assignment	Position (cm^{-1})				Description
	A_3T_3	$(AT)_3$	AT[a]	GC[b]	
AT_{2S}	1692	1689	1692		T_2 C=O stretch
G_SC_S(-)				1684	GC C=O antisymmetric stretch
AT_{4S}	1666	1666	1664		Base-paired T_4 C=O stretch
G_SC_S(+)				1651	GC C=O symmetric stretch
T_R	1646	1650	1640		T ring vibration
$A_{R1}T$	1622	1622	1622 (s)		Coupled AT ring vibration/
G_SC_R				1622 (w)	C ring mode + G C=O

[a] *obtained from the sequence: 5'-ATTATTATTATATTA-3' (Fig S1(a));* [b] *obtained from the sequence: 5'-GCCGCCGCCG-3' (Fig S1(b)); modes marked (s) are those which contribute strongly to the overall spectrum and modes marked (w) are those which contribute weakly to the overall spectrum.*

The shift of the AT_{2S} peak at 1690 cm^{-1} to higher frequency (blue arrow), as observed for the A_3T_3 sequence, is still detected, but the other features are less apparent. A contribution from a loss of intensity of the AT_{2S} mode means that the negative portion of this feature is of slightly larger magnitude than the positive component.

Despite the modest impact on the IR spectra, ligand-induced stabilization of the melting temperatures of the DNA duplexes by 24 °C and 16 °C was observed for the A_3T_3 and $(AT)_3$ sequences respectively, confirming H33258 binding (Fig.S3-S8). The reduced stabilization for the alternating sequence is consistent with the reported 22-fold preference of H33258 for binding to A-tracts.[8]

The 2D-IR spectrum of the H-A_3T_3 complex is shown in Fig.1(b). Each peak observed in the FTIR spectrum gives rise to a negative feature (red) located on the 2D-IR spectrum diagonal. These are assigned to the respective $\upsilon = 0 \longrightarrow 1$ transitions, each with an accompanying, positive (blue), $\upsilon = 1 \longrightarrow 2$ peak shifted by ~10 cm^{-1} to lower probe frequency. [33,36,39-42]

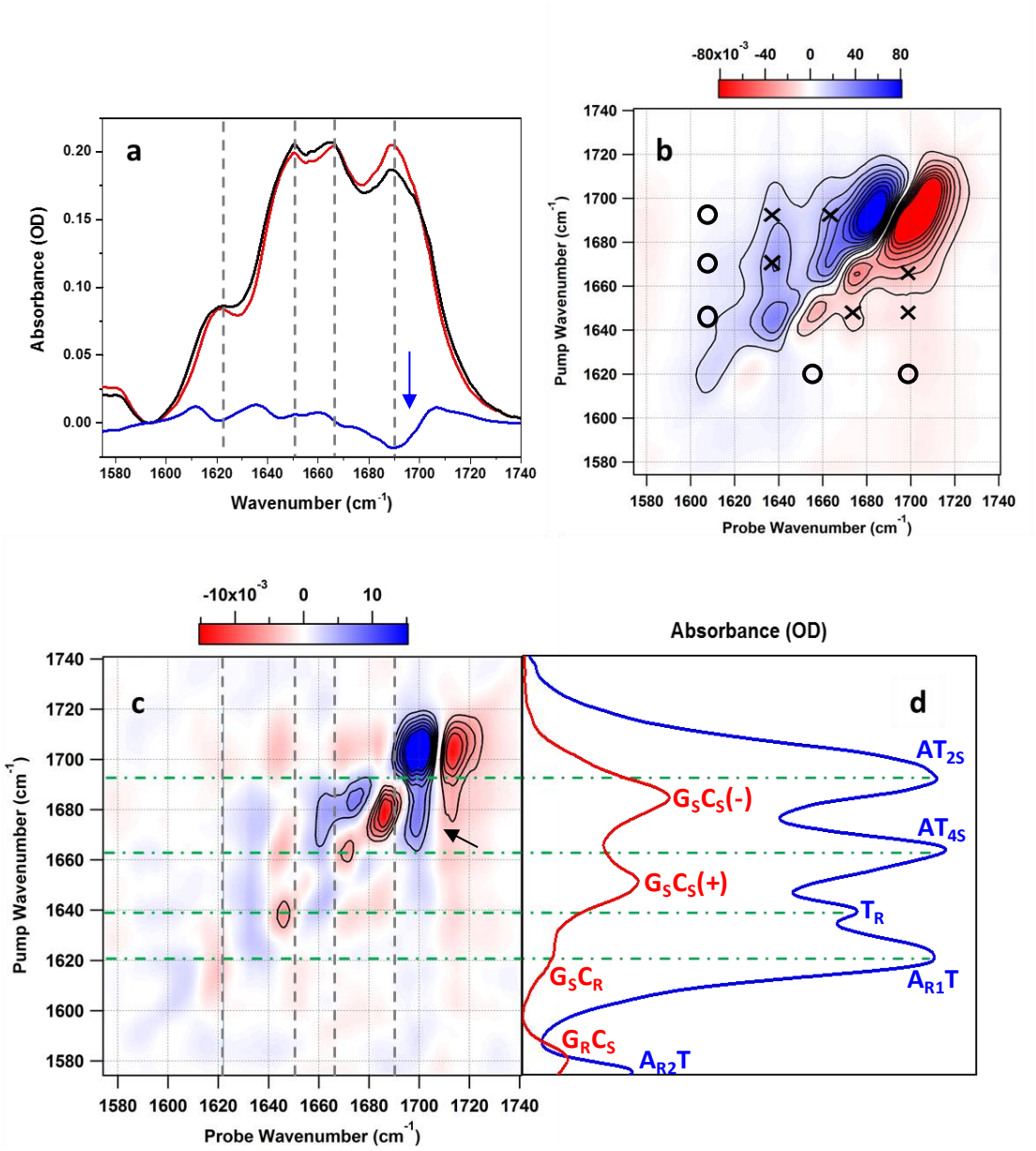

Figure 2: *a) FTIR spectra of $(AT)_3$ DNA with (black) and without (red) H33258. Blue spectrum shows the binding-induced difference FTIR spectrum (complex-free sequence). b) 2D-IR spectrum of H-$(AT)_3$ complex. Crosses mark off-diagonal peaks assigned to coupling of modes primarily located on T-base. Circles show off-diagonal peaks assigned to coupling of modes on A and T bases induced by W-C base pairing. c) Ligand binding induced 2D-IR difference spectrum. d) FTIR spectra of GC and AT-only DNA sequences to show the relative (weighted per base) magnitudes of peaks and peak positions to aid assignment of the IR spectrum of the $(AT)_3$ sequence.*

Peaks located in the off-diagonal region primarily indicate the presence of coupling between diagonal vibrational modes, though a small contribution is expected from energy transfer due to the fast (~650 fs) vibrational relaxation of the base vibrational modes.[33] Two groups of off-diagonal peaks are indicated in Fig.1(b). Those due to intra-base coupling between the AT_{2S}, AT_{4S} and T_R modes that are primarily located on the thymine base are marked with crosses. Other off-diagonal peaks (Fig 1(b), circles) indicate inter-base coupling between the adenine-based $A_{R1}T$ mode and the three thymine modes that is induced by Watson-Crick H-bonding.

A difference 2D-IR spectrum shows the impact of binding H33258 (Fig.1(c)). A comparison of the diagonal peak positions in the 2D-IR difference spectrum with the linear AT and GC spectra (Fig.1(d)) suggests that H33258 binding primarily affects the AT_{2S} and AT_{4S} vibrational modes of the A_3T_3 duplex. Other, smaller, features appear in the 2D-IR difference spectrum that originate from changes in the $A_{R1}T$ modes, consistent with the FTIR spectrum, but also the T_R mode and related off-diagonal peaks (crosses Fig.1(c)).

The 2D-IR spectrum of the H-$(AT)_3$ complex and the binding-induced 2D-IR difference spectrum for the $(AT)_3$ sequence are shown in Figs.2(b) and (c) respectively. The major changes upon binding affect the AT_{2S} and AT_{4S} modes and the overall pattern of peaks in the 2D-IR difference spectrum is similar to that for the A-tract sequence, although the AT_{2S} feature is clearly dominant.

The overlapping peaks in the IR spectrum of the DNA bases mean that the 2D-IR off-diagonal peaks are essential for clear differentiation between changes affecting the AT and GC base pairs (e.g. arrows Figs 1(c)&2(c)). The peaks in the difference FT-IR and 2D-IR spectra focus on the AT_{2S} and AT_{4S} modes, but to rule out contributions from GC modes and to quantify the changes, we employed a model 2D-IR spectrum (Fig.3) constructed from 2D Gaussian

Table 2: Changes to the IR response of A_3T_3 and $(AT)_3$ DNA duplexes due to H33258 binding.

Changes due to H33258 Binding	Host DNA Duplex	
	A_3T_3	**$(AT)_3$**
AT_{2S} Shift (cm^{-1})	6.4	15.1
AT_{2S} Shifted Subset Size (%)	**20.3**	**11.3**
Change in T_2 H-bond Strength ($kJmol^{-1}$)[a]	-2.8	-6.6
AT_{4S} Shift (cm^{-1})	8.7	15.6
AT_{4S} Shifted Subset Size (%)	**9.6**	**1.7**
Change in T_4 H-bond Strength ($kJmol^{-1}$)[a]	-3.9	-6.9

[a]Values calculated using the C=O bond energy (~ 743 $kJmol^{-1}$)[43] via the methodology as outlined in ref. 44 using equation ref. 45.

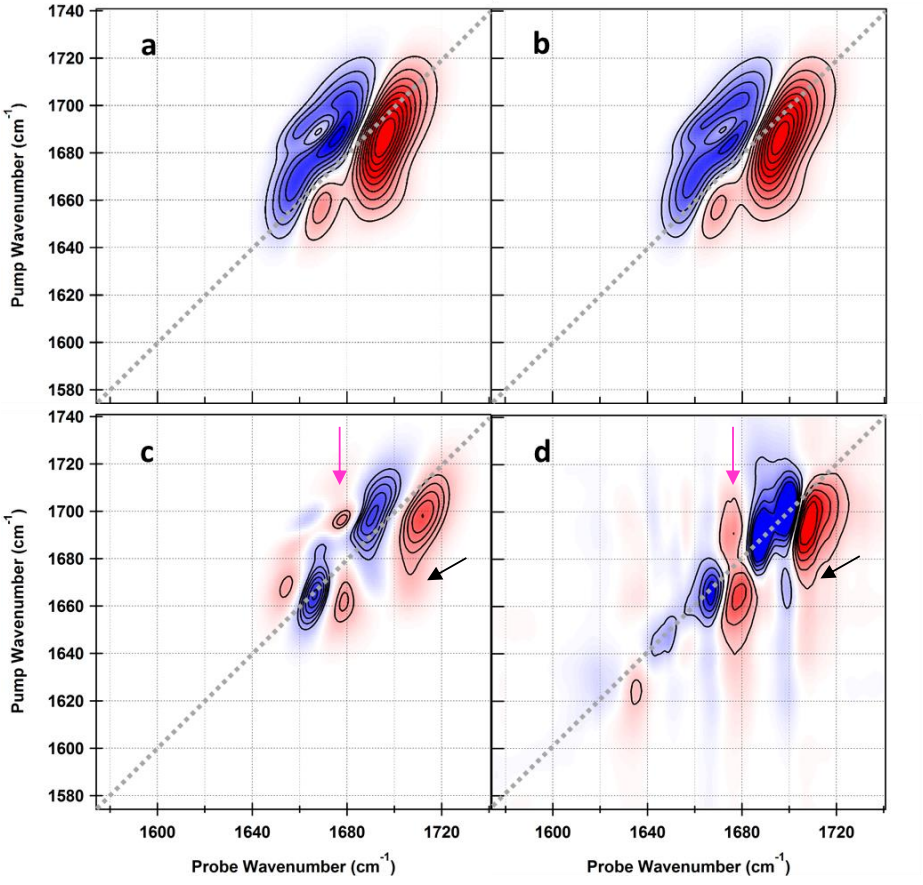

Figure 3: *Simulated 2D-IR spectra of the coupled AT_{2S} and AT_{4S} transitions before (a) and after (b) a shift of a subset of the bands to higher frequency. (c) shows the simulated difference 2D-IR spectrum that results from a subtraction of (a) from (b). (d) Shows the difference 2D-IR spectrum obtained following formation of the H-A_3T_3 complex. Pink and black arrows point to off-diagonal features linking the two modes in experimental and simulated spectra.*

lineshapes, which simulate the coupled peaks in the 2D-IR spectra of the DNA duplexes (Fig.3(a) and 3(b)). A variety of spectral changes were modelled and the difference spectra simulated (SI (Fig.S9)). The best agreement with the experimental 2D-IR difference spectra (Figs.1(c)&2(c)) was obtained by shifting a subset of the AT_{2S} and AT_{4S} peaks to higher frequency (Fig.3(c)). Crucially, this simulation recreated the off-diagonal structure of the difference spectrum as well as the diagonal peaks (see arrows Fig.3(c&d)). The results are summarized in Table 2, which show that ~20% of the AT_{2S} mode of the A_3T_3 (A-tract) sequence is shifted ~6 cm^{-1} to higher frequency while ~10% of the AT_{4S} mode shifts by ~9 cm^{-1}. For the alternating sequence, the shift to higher frequency is ~7-9 cm^{-1} larger for both modes, but the fraction of the carbonyls affected is smaller. Particularly noteworthy is the very small (2%) fraction of the AT_{4S} peak that is shifted whereas the 11% change in the AT_{2S} mode is more comparable with the A_3T_3 sequence.

Based on these models, we assign the peaks appearing in the FT-IR and 2D-IR difference spectra to a binding-induced shift to higher frequency of a portion of the AT_{2S} and AT_{4S} base vibrations. No contributions from diagonal or off-diagonal peaks attributable to GC were observed, showing that the binding interactions of H33258 are localized to the AT region of the duplex.

Shifts of the AT_{2S} and AT_{4S} modes of the A-tract sequence to higher frequency are consistent with a decrease in the strength of H-bonds formed to the T_2 and T_4 carbonyl groups upon H33258 binding of 2.8 and 3.9 $kJmol^{-1}$ respectively. [43-45] It is important to note that, although the two modes exhibit coupling, it has been demonstrated that this is weak and that a local mode picture, treating the two carbonyls as separate modes is appropriate. [29] The smaller changes in the $A_{R1}T$ and T_R vibrational modes are likely to be a result of changes to the two T carbonyls since their stretches are known to contribute to these modes.[39]

The T_2 carbonyls point directly into the minor groove of the DNA duplex and do not engage in W-C H-bonding. Water molecules are present in the minor groove as a spine of hydration and form H-bonds to the T_2 carbonyl. [10,11] The reduction in strength of these H-bonds to some of the T bases in the sequence is consistent with structural studies showing that the bis-benzimidazole of H33258 displaces water molecules from the minor groove. [10,11]

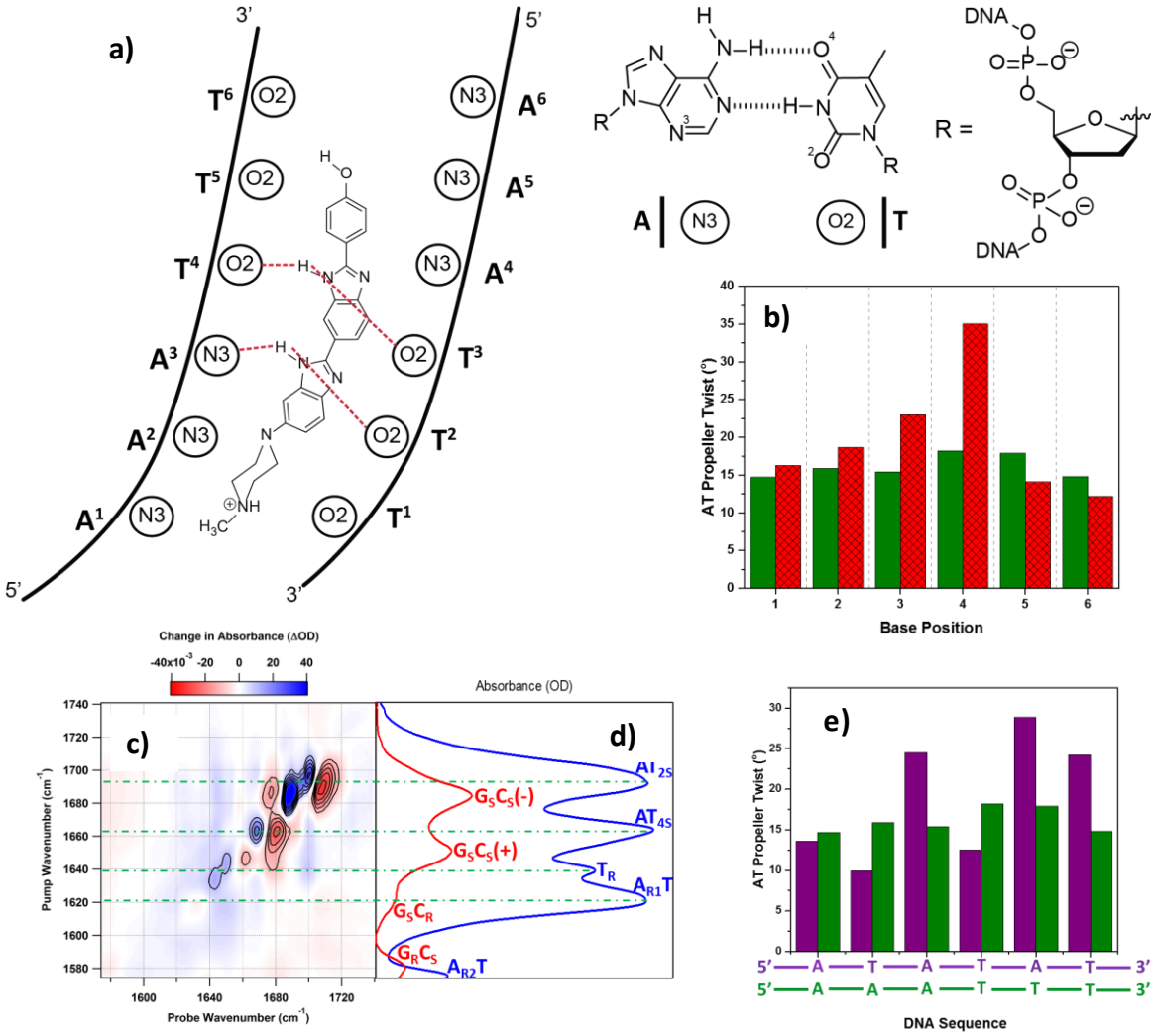

Figure 4: *a) Diagram illustrating the binding of H-A_3T_3 from X-ray crystallography.[18] Pink lines show H-bonds between ligand and DNA. b) AT propeller twists of the free (green bars, PDB ID: 1S2R[12]) and the complexed A_3T_3 (red bars, PDB ID: 296D[18]). c) 2D-IR difference spectrum between $(AT)_3$ and A_3T_3 duplex d) FTIR spectra of AT only (blue) and GC only (red) DNA sequences e) AT propeller twists of A_3T_3 (green, PDB ID: 1S2R[12]) and $(AT)_3$ (purple, PDB ID: 1DN9[13]) base sequences.*

Although the ligand has been shown to replace the H-bonds to the T_2 carbonyls, our results show that the overall H-bond strength is lower than that formed by the water, consistent with evidence that the interaction between H33258 and DNA is not stabilized to any great degree by H-bonding.[11]

The frequency shift of the T_4 carbonyl, which is involved in the W-C interaction, shows that binding is also accompanied by a structural change affecting the DNA bases. The two major structural changes that can weaken the W-C hydrogen bond are a buckle and propeller twist.[40] An increase in the base buckle angle will reduce the strength of the W-C H-bonds, but a computational study predicted that the T_4 carbonyl is not strongly affected by this until an extreme angle of buckle (~50°) is reached. [40] This is not consistent with the relatively minor structural changes detected by NMR[20-23] or crystallography[15-19] on H33258 binding to the A_3T_3 duplex.[15] An increase in the propeller twist will cause an increase in the length of the W-C hydrogen bond to the T_4 carbonyl, decreasing its strength. This is consistent both with the spectroscopic data and previous structural studies using 2D-NMR[20-23] and X-ray crystallography,[15-19] which have reported changes in propeller twist. The results of crystallographic investigations of the structure of the H-A_3T_3 complex are reproduced in (Fig.4(a)) alongside a bar-graph showing perturbations in the propeller twist in the H-A_3T_3 complex (Fig.4(b), red bars) relative to the uncomplexed DNA (Fig.4(b), green). H33258 forms two bifurcated H-bonds to T_2 carbonyls (Fig.4(a), pink lines) and accommodating these interactions increases the propeller twist in some of the AT base pairs with one (at position 4) particularly strongly affected. This is consistent with our observations of a subset of T_4-related modes shifting to higher frequency upon ligand binding. The propeller twisting does not affect the GC base pairs, again consistent with our spectroscopic data. Whilst supporting the conclusions of crystallographic studies, our new results demonstrate that, in solution, the influence of solvated water does not perturb the nature of the interaction. Whilst at first this is

expected given the nature of binding and rejection of water within the minor grove, we show that, even in solution, the interaction is localised by not observing spectral perturbations in the neighboring GC.

Comparing the A-tract sequence with the alternating sequence, it is clear that the disruption to the spine of hydration, affecting the T_2 carbonyl is still present in the data. The smaller shifted subset size and the larger frequency shift observed for the T_2 carbonyl in the $(AT)_3$ sequence supports the conclusion of an X-ray crystallography[19] study of Hoechst33258 binding to a $(AT)_2$ sequence, which indicated the presence of only one bifurcated H-bond as opposed to two in the complex with A_3T_3. However, the impact upon the propeller twist, while present, is markedly reduced. This is shown by the relatively small size (1.7%) of the shifted subset in the $(AT)_3$ duplex showing that this does not undergo significant structural changes upon binding compared to the A_3T_3 (9.7%) sequence. This is a major difference between the two strands and may be relevant to the differences in binding affinity. It is clear from thermodynamic measurements that H33258 binding is entropically-driven; this is attributed to the release of water from the spine of hydration. Our data reflects this and the minor groove of the A-tract sequence is known to contain a more ordered spine of hydration than the alternating sequence. [12-14] Many reports assume that conformational changes in the DNA accompanying this binding are negligible, citing a 'lock and key' or 'rigid body' interaction. [10,11,46-48] In contrast, our data indicate that conformational changes take place in order to accommodate and optimize structures for base-to-H33258 binding, which would be expected to increase favorable enthalpic contributions to binding, as given by the induced fit model.

The enthalpy of binding has been measured to be endothermic overall, but the greatest favourable contribution to binding has been assigned to hydrophobic contacts between the DNA groove and the ligand. [11,48] A-tract DNA is unique in that it features a very regular set of propeller twists (Fig 4(b) green bars) due to the formation of 3-centre H-bonds, leading to a

rigid and ordered section of the helix. [43] The indications from our data are that, rather than this unique structural feature being beneficial to ligand binding in itself, the disruption of the A-tract structure leads to a closer interaction between ligand and DNA than is possible in the alternating sequence. While the change in macroscopic conformation of the DNA is small and unlikely to outweigh the entropic benefits of releasing hydration water, our results indicate that binding to A-tract DNA is better described by an induced fit type model, where the DNA structure changes to accommodate the ligand. By contrast, binding to a sub-optimal sequence ($(AT)_3$) proceeds without significant change to the DNA structure, which lacks the enthalpic benefit of the improved DNA-ligand interactions and also does not gain from any entropic benefit from disrupting the ordered propeller twists. A structure-related narrowing of the minor groove and restriction of dynamical motion upon ligand binding to A-tract DNA has been suggested in simulations but was stated to be insignificant.[43,49] Rather, our results support conclusions drawn from NMR studies, demonstrating that induced fit is an important aspect of optimized Hoechst binding.[24]

Further evidence that the difference 2D-IR spectrum observed on the formation of the H-A_3T_3 complex results from an order-disorder change in the base orientation within the AT region of the duplex arises from a difference 2D-IR spectrum between the uncomplexed A_3T_3 and $(AT)_3$ sequences (Fig.4.(c)). These structures differ mainly through the three-centered H-bonds and ordered propeller twists of the A_3T_3 sequence (Fig.4(e), green) in contrast to the alternating sequence (Fig.4.(e), purple).[13] The difference 2D-IR spectrum (Fig.4.(c)) is remarkably similar to that arising from formation of the H-A_3T_3 complex; the AT_{2S} and AT_{4S} modes shift by 6.3 cm^{-1} and 10.7 cm^{-1}. This observation supports the hypothesis that the ligand binding induces changes in the propeller twists of the A_3T_3 sequence. It is also noticeable that the T_2 mode shifts between the two uncomplexed sequences. This suggests a weakening of the hydrogen bonds to the T_2 in the alternating sequence versus the A-tract, consistent with the more ordered spine of

hydration (and so stronger H-bonds) found in the minor groove of the latter sequence. [12-14] This further indicates that an entropic gain may arise from losing the spine of hydration as well as the ordered propeller twist caused by induced fit of the ligand.

Conclusions:

In conclusion, 2D-IR shows that binding H33258 to two DNA sequences leads to shifts in vibrational modes associated with specific AT base pairs that are due to the loss of the spine of hydration and formation of direct H-bonding between DNA and ligand as well as alterations in the propeller twist induced by the ligand locating in the minor groove. Comparison of binding to A-tract and alternating DNA sequences revealed that binding to A_3T_3 results in loss of the ordered propeller twist arrangement of bases found in the uncomplexed DNA. This is not replicated to the same extent in the alternating sequence and we propose that these structural changes constitute an induced fit type interaction that facilitates superior accommodation of H33258 and increased hydrophobic interactions between ligand and DNA. This contradicts current pictures which treat H33258 binding as a rigid body interaction and complements the entropic release of water from the minor groove. Finally, the results fully demonstrate 2D-IR capabilities to simplify quantification of solution phase DNA-binding.

Supporting Information

Additional information relating to experimental details, spectroscopic characterization of samples and 2D-IR data analysis can be found in the Supporting Information.

Acknowledgments

The authors wish to acknowledge funding from the Science and Technologies Facilities Council (STFC) for Programme Access to the ULTRA spectrometer at the Central Laser Facility.

References

(1) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids, *Nature*, **1953**, *4356*, 737.

(2) Neidle, S. DNA Minor-Groove Recognition by Small Molecules. *Nat. Prod. Rep.* **2001**, *18*, 291-309.

(3) White, C. M.; Heidenreich, O.; Nordheim, A.; Beerman, T. A. Evaluation of the Effectiveness of DNA-Binding Drugs to Inhibit Transcription Using the c-fos Serum Response Element as a Target, *Biochemistry*, **2000**, *39*, 12262-12273

(4) Zhang, X.; Kiechle, F. Hoechst33342-Induced Apoptosis is Associated with Decreased Immunoreactive Topoisomerase I and Topoisomerase I-DNA Complex Formation, *Annals of Clinical & Laboratory Science*, **2001**, *31*, 187-198

(5) Bellorini, M.; Moncollin, V.; D'Incalci, M.; Mongelli, N.; Mantovani, R. Distamycin A and Tallimustine Inhibit TBP Binding and Basal *In Vitro* Transcription, *Nucleic Acids Research*, **1995**, *23*, 1657-1663.

(6) Harshman, K. D.; Dervan, P. B. Molecular Recognition of B-DNA by Hoechst33258. *Nucleic Acids Research* **1985**, *13*, 4825-4835.

(7) Drobyshev, A. L.; Zasedatelev, A. S.; Yershov, G. M.; Mirzabekov, A. D. Massive Parallel Analysis of DNA-Hoechst33258 Binding Specificity with a Generic Oligodeoxyribonucleotide Microchip. *Nucleic Acids Research*, **1999**, *27*, 4100-4105.

(8) Breusegem, S. Y.; Clegg, R. M.; Loontiens, F. G. Base-Sequence Specificity of Hoechst33258 and DAPI Binding to Five $(A/T)_4$ DNA Sites with Kinetic Evidence for More Than One High-Affinity Hoechst33258-AATT Complex, *J. Mol. Biol.*, **2002**, *315*, 1049-1061.

(9) Abu-Daya, A.; Brown, P. M.; Fox, K. R. DNA Sequence Preferences of Several AT-Selective Minor Groove Binding Ligands, *Nucleic Acids Research*, **1995**, *23*, 3385-3392.

(10) Haq, I. Part II: The Thermodynamics of Drug-Bipolymer Interaction, Thermodynamics of Drug-DNA Interactions. *Archives of Biochemistry and Biophysics*, **2002**, *403*, 1-15.

(11) Haq, I.; Ladburry, J. E.; Chowdhry, B. Z.; Jenkins, T. C.; Chairs, J. B. Specific Binding of Hoechst33258 to the d(CGCAAATTTGCG)$_2$ Duplex: Calorimetric and Spectroscopic Studies. *J. Mol. Biol.* **1997**, *271*, 244-257.

(12) Woods, K. K.; Maehigashi, T.; Howerton, S. B.; Sines, C. C.; Tannenbaum, S. T.; Williams, L. D.; High-Resolution Structure of an Extended A-Tract: [d(CGCAAATTTGCG)]$_2$, *J. Am. Chem. Soc.*, **2004**, *126*, 15330-15331.

(13) Yoon, C.; Privé, G. G.; Goodsell, D. S.; Dickerson, R. E. Structure of an Alternating-B DNA Helix and its Relationship to A-tract DNA, *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 6332-6336.

(14) Edwards, K. J.; Brown, D. G.; Spink, N.; Neidle, S. Molecular Structure of the B-DNA Dodecamer d(CGCAAATTTGCG)2. An examination of Propeller Twist and Minor-Groove Water Structure at 2.2 A Resolution, *J. Mol. Biol.*, **1992**, *226*, 1161-1173.

(15) Teng, M.; Usman, N.; Frederick, C. A., Wang, A. H.-J. The Molecular Structure of the Complex of Hoechst33258 and the DNA Dodecamer d(CGCGAATTCGCG). *Nucleic Acids Research*, **1988**, *16*, 2671-2690

(16) Quintana, J. R.; Lipanov, A. A.; Dickerson, R. E. Low-Temperature Crystallographic Analyses pf the Binding of Hoechst33258 to the Double-Helical DNA Dodecamer C-G-C-G-A-A-T-T-C-G-C-G, *Biochemistry*, **1991**, *30*, 10294-10306

(17) Vega, M. C.; Sáez, I. G.; Aymami, J.; Erita, R.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Coll, M. Three-Dimensional Crystal Structure of the A-tract DNA Dodecamer d(CGCAAATTTGCG) Complexed with the Minor-Groove-Binding Drug Hoechst33258, *Eur. J. Biochem.*, **1994**, *222*, 721-726.

(18) Spink, N.; Brown, D. G.; Skelly J. V.; Neidle, S. Sequence-Dependent Effects in Drug-DNA Interaction: The Crystal Structure of Hoechst33258 Bound to the d(CGCAAATTTGCG)$_2$ Duplex, *Nucleic Acids Research*, **1994**, *22*, 1607-1612.

(19) Carrondo, M. A. A. F.; De C. T.; Coll, M.; Aymami, J.; Wang, A. H.-J.; van der Marel, G. A.; van Boom, J. H.; Rich, A. Binding of Hoechst Dye to d(CGCGATATCGCG) and its Influence on the Conformation of the DNA Fragment. *Biochemistry*, **1989**, *28*, 7849-7859.

(20) Parkinson, J. A.; Barber, J.; Douglas, K. T.; Rosamond, J.; Sharples, D. Minor-Groove Recognition of the Self-Complementary Duplex $d(CGCGAATTCGCG)_2$ by Hoechst33258: A High-Field NMR Study. *Biochemistry*, **1990**, *29*, 10181-10190.

(21) Fede, A.; Labhardt, A.; Bannwarth, W.; Leupin, W. Dynamics and Binding Mode of Hoechst33258 to $d(GTGGAATTCCAC)_2$ in the 1:1 Solution Complex as Determined by Two-Dimensional ^1H NMR. *Biochemistry*, **1991**, *30*, 11377-11388

(22) Higgins, L. D.; Searle, M. S. Site-Specificity of Bis-Benzimidazole Hoechst33258 in A-tract Recognition of the DNA Dodecamer Duplex $d(GCAAAATTTTGC)_2$. *Chem. Commun.,* **1999**, *18*, 1861-1862.

(23) Parkinson, J. A.; Ebrahimi, S. E.; McKie, J. H.; Douglas, K. T. Molecular Design of DNA-Directed Ligands with Specific Interactions: Solution NMR Studies of the Interaction of a m-Hydroxy Analog of Hoechst33258 with d(CGCGAATTCGCG)2, *Biochemistry*, **1994**, *33*, 8442-8452.

(24) Bostock-Smith, C. E.; Harris, S. A.; Laughton, C. A.; Searle, M. S. Induced Fit DNA Recognition by a Minor Groove Binding Analogue of Hoechst33258: Fluctuations in DNA A-tract Structure Investigated by NMR and Molecular Dynamics Simulations, *Nucleic Acids Research*, **2001**, *29*, 693-702.

(25) Hunt, N. T. 2D-IR spectroscopy: Ultrafast Insights into Biomolecule Structure and Function, *Chem. Soc. Rev.*, **2009**, *38*, 1837-1848.

(26) Park, S.; Kwak, K.; Fayer, M. D. Ultrafast 2D-IR Vibrational Echo Spectroscopy: A Probe of Molecular Dynamics, *Laser Phys. Lett.*, **2007**, *10*, 704-718.

(27) Baiz, C. R.; Mcrobbie, P. L.; Anna, J. M.; Geva, E.; Kubarych, K. J. Two-Dimensional Infrared Spectroscopy of Metal Carbonyls, *Accounts of Chemical Research*, **2009**, *42*, 1395-1404.

(28) Downs, T. R.; Wilfinger, W. W. Fluorometric Qualification of DNA in Cells and Tissue. *Anal. Biochem.* **1983**, *131*, 538-547.

(29) Peng, C. S.; Jones, K. C.; Tokmakoff, A. Anharmonic Vibrational Modes of Nucleic Acid Bases Revealed by 2D IR Spectroscopy, *J. Am. Chem. Soc.*, **2011**, *133*, 15650-15660.

(30) Yang, M.; Szyc, L.; Elsaesser, T. Femtosecond Two-Dimensional Infrared Spectroscopy of Adenine-Thymine Base Pairs in DNA Oligomers, *J. Phys. Chem. B*, **2011**, *115*, 1262-1267.

(31) Greve, C.; Elsaesser, T. Ultrafast Two-Dimensional Infrared Spectroscopy of Guanine-Cytosine Base Pairs in DNA Oligomers, *J. Phys. Chem. B*, **2013**, *117*, 14009-14017.

(32) Krummel, A. T.; Zanni, M. T. DNA Vibrational Coupling Revealed with Two-Dimensional Infrared Spectroscopy: Insight into Why Vibrational Spectroscopy is Sensitive to DNA Structure, *J. Phys. Chem. B*, **2006**, *110*, 13991-14000.

(33) Hithell, G.; Shaw, D. J.; Donaldson, P. M.; Greetham, G. M.; Towrie, M.; Burley, G. A.; Parker, W. A.; Hunt, N. T. Long-Range Vibrational Dynamics Are Directed by Watson-Crick Base Pairing in Duplex DNA. *J. Phys. Chem. B*, **2016**, *120*, 4009-4018.

(34) Guchhait, B.; Liu, Y.; Siebert, T.; Elsaesser, T. Ultrafast Vibrational Dynamics of the DNA Backbone at Different Hydration Levels Mapped by Two-Dimensional Infrared Spectroscopy, *Structural Dynamics*, **2016**, *3*, 043202-043217.

(35) Yang. M.; Szyc, L.; Elsaesser, T. Decelerated Water Dynamics and Vibrational Couplings of Hydrated DNA Mapped by Two-Dimensional Infrared Spectroscopy, *J. Phys. Chem. B*, **2011**, *115*, 13093-13100.

(36) Sanstead, P. J.; Stevenson, P.; Tokmakoff, A. Sequence-Dependent Mechanism of DNA Oligonucleotide Dehybridization Resolved through Infrared Spectroscopy. *J. Am. Chem. Soc.*, **2016**, *138*, 11792-11801.

(37) Shaw, D. J.; Adamczyk, K., Frederix, P. W. J. M.; Simpson, N.; Robb, K.; Greetham, G. M.; Towrie, M.; Parker, A. W.; Hoskisson, P. A.; Hunt, N. T. Multidimensional Infrared Spectroscopy Reveals the Vibrational and Solvational Dynamics of Isoniazid, *J. Chem. Phys*, **2015**, *142*, 212401.

(38) Greetham, G. M.; Burgos, P.; Cao, Q.; Clark, I. P.; Codd, P. S.; Farrow, R. C.; George, M. W.; Kogimtzis, M.; Matousek, P.; Parker, A. W.; *et al*. M. ULTRA: A Unique Instrument for Time-Resolved Spectroscopy. *Applied Spectroscopy*, **2010**, *12*, 1311-1319.

(39) Lee, C.; Cho, M. Vibrational Dynamics of DNA. II. Deuterium Exchange Effects and Simulated IR Absorption Spectra, *J. Chem. Phys.*, **2006**, *125*, 114509

(40) Lee, C.; Park, K.-H.; Cho, M. Vibrational Dynamics of DNA. I. Vibrational Basis Modes and Couplings, *J. Chem. Phys.*, **2006**, *125*, 114508

(41) Lee, C.; Park, K.-H.; Kim, J.-A.; Hahn, S.; Cho, M. Vibrational Dynamics of DNA. III. Molecular Dynamics Simulations of DNA in Water and Theoretical Calculations of the Two-Dimensional Vibrational Spectra, *J. Chem. Phys.*, **2006**, *125*, 114510

(42) Lee, C.; Cho, M. Vibrational Dynamics of DNA. IV. Vibrational Spectroscopic Characteristics of A-, B-, and Z-Form DNAs, *J. Chem. Phys.*, **2007**, *126*, 145102.

(43) Mukerji, I.; Williams, A. P. UV Resonance Raman and Circular Dichroism Studies of a DNA Duplex Containing an A_3T_3 Tract: Evidence for a Premelting Transition and Three-Centered H-Bonds, *Biochemistry*, **2002**, *41*, 69-77.

(44) Thorogood, H.; Waters, T. R.; Parker, A. W.; Wharton, C. W.; Connolly, B. A. Resonance Raman Spectroscopy of 4-Thiothymidine and Oligodeoxynucleotides Containing This Base Both Free in Solution and Bound to the Restriction Endonuclease *Eco*RV. *Biochemistry*, **1996**, *35*, 8723-8733.

(45) $\Delta\Delta H = \left(\frac{\Delta\nu}{\upsilon}\right) U_{Bond}$

where $\Delta\Delta H$ is the change in the hydrogen bond strength, $\Delta\nu$ is the shift in the frequency of the vibrational mode, υ is the original frequency of the vibrational mode and U_{Bond}

is the bond energy of the affected moiety. Note this equation in only valid if $\Delta\nu$ is small compared to U_{Bond}.

(46) Kiser, J. R.; Monk, R. W.; Smalls R. L.; Petty, J. T. Hydration Changes in the Association of Hoechst33258 with DNA. *Biochemistry*, **2005**, *44*, 16988-16997.

(47) Fresch, B.; Remacle, F. Atomistic Account of Structural and Dynamical Changes Induced by Small Binders in the Double Helix of a Short DNA, *Phys. Chem. Chem. Phys.*, **2014**, *16*, 14070-14082.

(48) Han, F.; Taulier, N.; Chalikian, T. V. Association of the Minor Groove Binding Drug Hoechst33258 with d(CGCGAATTCGCG)$_2$: Volumetric, Calorimetric, and Spectroscopic Characterizations, *Biochemistry*, **2005**, *44*, 9785-9794.

(49) Movileanu, L.; Benevides, J. M.; Thomas Jr, G. J. Determination of Base and Backbone Contributions to the Thermodynamics of Premelting and Melting Transitions in *B* DNA, *Nucleic Acids Research*, **2002**, *30*, 3797-3777.

TOC Graphic