COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Drug-induced stabilisation of a mismatched C-T base pair in a DNA hairpin†

Cathal T. Gallagher and Mark S. Searle*
School of Chemistry, University Park, Nottingham, UK NG7 2RD. E-mail: mark.searle@nottingham.ac.uk

Received (in Cambridge, UK) 12th May 2003, Accepted 17th June 2003
First published as an Advance Article on the web 1st July 2003

We have shown that a key feature of drug binding, namely specific G-C base pair recognition at a 5′-TG step, can induce a number of novel structural features when an extrahelical base is inserted in close proximity to the drug binding site; we have clearly demonstrated the formation of a stabilised C-T mismatched base pair at a non-terminal site.

Nucleic acids can adopt a wide range of sequence-specific conformations stabilised by weak non-covalent forces. Unpaired or bulged bases can be accommodated into an otherwise complementary DNA double helical conformation either through intrahelical stacking or by flipping the base out of the stack into solution.[1] Alternatively, a DNA frameshift can occur with the potential incorporation of a mismatched base pair. Single base bulges have been of particular interest because DNA mutations can arise due to DNA strand misalignment during replication.[2] Whether these mutations result in an addition of a base or a deletion appears to depend on the conformational preference of the extra base. The conformation and dynamics of this extra base may be strongly influenced by the base-specific interaction of drug molecules. Binding close to such a site has the potential to stabilise non-Watson–Crick base pairs or base bulges, increasing the susceptibility to mutations.[3,4]

We have investigated the effects of drug binding on DNA conformation and strand alignment using the anthracycline antibiotic nogalamycin (Fig. 1a) as a highly sequence specific probe binding preferentially to 5′-TG sites. Nogalamycin is a threading intercalator that binds to duplex DNA with bulky sugar residues located in both grooves of the double helix. Structural studies have shown that interactions in the major groove give the drug a unique binding orientation dictated by the G base.[5] We report NMR studies of a DNA hairpin sequence with a double-stranded stem region containing a single high affinity 5′-TG binding site with an adjacent mispaired T (see $H1^m$ in Fig. 1b).

We have employed a DNA hairpin motif to offset the inherent destabilising effects of introducing non-Watson–Crick base pairs in a short duplex. We have used the 5′-GNA loop sequence (N = G, A, C or T) that has previously been reported to show remarkable thermal stability in very short mini-hairpins;[6] the 5′-GTA loop gives superior chemical shift dispersion for structural analysis. The hairpin sequence (H1) is that shown in Fig. 1b.[7] The conformation of the hairpin in the absence of drug shows characteristic NMR parameters that demonstrate that the 5′-GTA loop is formed.[6] Examination of the imino proton region of the spectrum at 278 K reveals only one stabilised Watson–Crick base pair corresponding to the G immediately adjacent to the G-A pair in the loop. Detailed NMR analysis shows that while the stem region close to the loop is formed the ends of the hairpin are destabilised by the introduction of the extra T and are largely unstructured (H1 in Fig. 1b).

Complex formation with nogalamycin was monitored by 1D ^{1}H NMR titration studies and showed that the free and bound DNA species were in slow exchange on the chemical shift timescale. The resonances of the free DNA are replaced not by a single set of new resonances but by two sets corresponding to two species that are approximately equally populated at 278 K. Given the requirement at the binding site for a 5′-TG step, we can envisage two possible drug-induced DNA conformations that are consistent with the data (Fig. 1b, $H1^b$ and $H1^m$). The first ($H1^b$) has the extra T flipped out of the stack resulting in a one base frameshift preserving Watson–Crick base pairing down stream but resulting in a single C overhang. In the second case ($H1^m$), the drug stabilises a mismatched C-T basepair at the drug intercalation site.

In previous studies the stacking of an A-T base pair directly on the drug results in an upfield shift of the thymine N3-H by ~1 ppm to ~13 ppm.[5] However, we clearly see a resolved peak at 14.0 ppm in a position characteristic of an unperturbed Watson–Crick A-T base pair. Since H1 has only a single A-T base pair this resonance can be attributed to structure $H1^m$, with a C-T mismatched pair on one side of the intercalation site. In variable-temperature studies, we see that the signal at 14.0 ppm becomes exchange broadened above 298 K with evidence for rapid interconversion between conformers (see ESI).† The $H1^b$

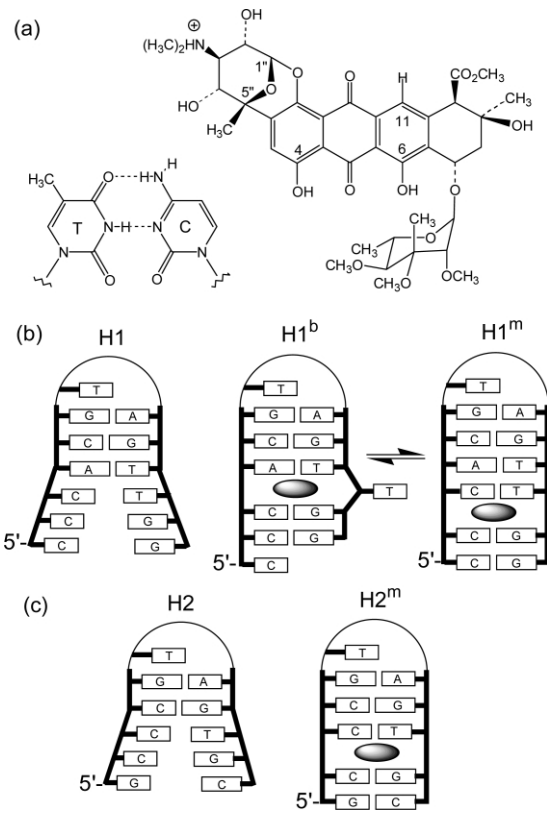

Fig. 1 (a) Structure of nogalamycin (Ng) and a C-T mismatch base pair; (b) sequence and folding of hairpin H1 with bulged ($H1^b$) and mismatch ($H1^m$) structures formed in the presence of drug; (c) hairpin H2 and structure of the mismatch complex with drug ($H2^m$).

† Electronic supplementary information (ESI) available: ^{1}H NMR spectra of the hairpin sequence. See http://www.rsc.org/suppdata/cc/b3/b305337g/

DOI: 10.1039/b305337g

conformer is more difficult to characterise due to line broadening effects and resonance overlap that preclude a detailed NMR analysis of the mixture of conformers. However, we have previously reported a detailed structural analysis of a Ng-hairpin complex in which an extra T inserted in the complementary strand 5′-C^{T}A is flipped out into solution in an analogous fashion to H1^{b}.[8]

In the current context, the C-T mismatch complex (H1^{m}) was of particular interest. To characterise this further we redesigned our hairpin sequence to eliminate the possibility of forming multiple species in solution. Thus, hairpin H2 (Fig. 1c) can give rise to only one possible complex containing a high affinity nogalamycin 5′-TG binding site. Again, studies of the free hairpin show that the 5′-GTA loop is formed along with the flanking C-G pair, however, the C-T mismatch destabilises the ends of the stem region. Drug binding to H2 produces a clean 1 : 1 complex with only a single species present in solution that proves amenable to a detailed NMR analysis. 2D NMR data collected at 278 and 288 K have enabled us to obtain a complete sequential assignment and identify intermolecular NOEs and chemical shift changes consistent with the proposed structure of H2^{m} (Fig. 1c).

The sequential connectivities are traced out in Fig. 2; the data show that the NOEs are contiguous except where weak or interrupted due to structural distortions across the 5′-TG intercalation site and close to the hairpin loop. NOEs from Ng-H11 to the deoxyribose H1′ of the flanking nucleotides T9 and G10 (peaks A and B in Fig. 2) unambiguously establish binding at the 5′-TG step. The stacking of the T-base results in an unusual upfield shift of the T-CH_3 to 0.64 ppm which is clearly indicative of direct stacking with the bound drug. Strong NOEs from the T-CH_3 to Ng 5″-CH_3 and Ng H1″ confirm this interaction.

A detailed analysis of the NOE data identifies many drug–DNA interactions in both grooves and has enabled us to calculate a family of structures of the Ng-H2^{m} complex using restrained molecular dynamics (MD) to model this unusual C-T mismatch (Fig. 3). Although intermolecular NOEs confirm the stacking of the C-T pair with the intercalated drug, the N3-H of the T base is exchange broadened suggesting that the C-T mismatch is not stabilised by strong base pair hydrogen bonding. This is confirmed by MD simulations that show that the average N3–N3 hydrogen bond distance is long (3.4 ± 0.4 Å). The formation of a stable C-T base pair requires a significant narrowing of the minor groove which, in this case, is partly precluded by steric interactions with the bound drug. Thus, stacking interactions rather than hydrogen bonding appear to stabilise the C-T pair.

We have shown that a key feature of drug binding, namely specific G-C base pair recognition at a 5′-TG step, can induce a

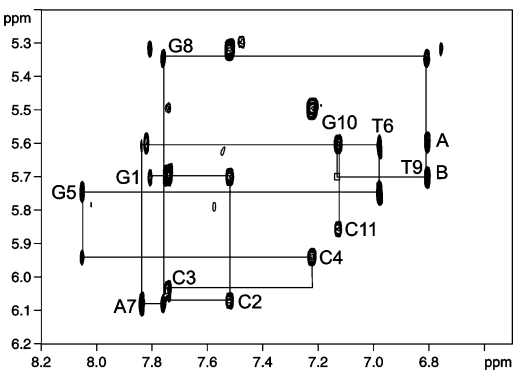

Fig. 2 600 MHz 2D NOESY data of the Ng-H2^{m} complex (288 K, pH 7.0) showing the sequential connectivity pathway for base H6/H8-deoxyribose H1′; peaks **A** and **B** are Ng-H11 to G10 and T9 H1′ [sequence numbering d(G1C2C3C4G5T6A7G8T9G10C11)]. Note that the chemical shifts of T9H6 and Ng-H11 are partially overlapped.

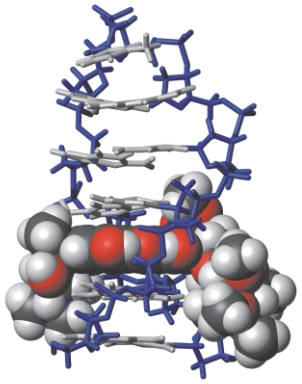

Fig. 3 Energy minimised structure of the nogalamycin H2^{m} complex from NOE restrained MD simulations at 300 K.

number of novel structural features when an extrahelical base is inserted in close proximity to the drug binding site. Nogalamycin binding to 5′-CG terminal bulge sites in two DNA heptamers, C^{T}GTACG and CGTAC^{T}G,[9] have reported a strand misalignment with stabilisation of C-T and G-T mispairs, leaving 5′-C and 3′-G dangling nucleotides, respectively. The structural distortions observed are likely to be influenced by end-effects in this context; however, here we have clearly demonstrated the formation of a stabilised C-T mismatched base pair at a non-terminal site. DNA intercalators have already been shown to have mutagenic effects during replication by perturbing the recognition of DNA defects by proofreading and repair enzymes.[3,10,11] In the current study we have shed some light on structural aspects of drug-DNA recognition in the context of inducing unusual base mis-pairing or bulged bases.

We thank the EPSRC for a studentship to C. T. G. and Dr H. Williams for guidance with molecular modelling. NMR facilities are supported by the EPSRC and the School of Chemistry.

Notes and references

1 S. A. Woodson and D. M. Crothers, *Biochemistry*, 1988, **27**, 3130; M. W. Kalnik, D. G. Norman, M. G. Zagorski, P. F. Swann and D. J. Patel, *Biochemistry*, 1989, **28**, 294.
2 G. Streisinger and J. Owen, *Genetics*, 1985, 633; S. A. Woodson and D. M. Crothers, *Biochemistry*, 1988, **27**, 436.
3 J. W. Nelson and I. Tinoco, *Biochemistry*, 1985, **24**, 6416; S. A. Woodson and D. M. Crothers, *Biochemistry*, 1988, **27**, 8904; K. Nakatani, S. Sando and I. Saito, *J. Am. Chem. Soc.*, 2000, **122**, 2172.
4 L. S. Kappen and I. H. Goldberg, *Biochemistry*, 1997, **36**, 14861; Z. Xi, Q. K. Mao and I. H. Goldberg, *Biochemistry*, 1999, **38**, 4342.
5 X. L. Zhang and D. J. Patel, *Biochemistry*, 1990, **29**, 9451; C. K. Smith, J. A. Brannigan and M. H. Moore, *J. Mol. Biol.*, 1996, **263**, 237; H. E. Williams and M. S. Searle, *J. Mol. Biol.*, 1999, **290**, 699.
6 I. Hirao, G. Kawai, S. Yoshizawa, Y. Nishimura, Y. Ishido, K. A. Watanabe and K.-I. Miura, *Nucleic Acids Res.*, 1994, **22**, 576; S. Yoshizawa, G. Kawai, K. A. Watanabe, K.-I. Miura and I. Hirao, *Biochemistry*, 1997, **36**, 4761; H. E. L. Williams, M. L. Colgrave and M. S. Searle, *Eur. J. Biochem.*, 2002, **269**, 1726.
7 Oligonucleotides were synthesised using standard solid-phase methods and purified by reverse-phase HPLC. Structures were calculated using procedures described in ref. 7 and were displayed and analysed using MOLMOL R. Koradi, M. Billeter and K. Wüthrich, *J. Mol. Graphics*, 1996, **14**, 51.
8 M. L. Colgrave, H. E. L. Williams and M. S. Searle, *Angew. Chem., Int. Ed.*, 2002, **41**, 4754.
9 J. Caceres-Cortes and A. H.-J. Wang, *Biochemistry*, 1996, **35**, 616.
10 W.-C. Lam, E. J. C. Van der Schans, L. C. Sowers and D. P. Millar, *Biochemistry*, 1999, **38**, 2661; V. A. Malkov, I. Biswas, R. D. Camerini-Otero and P. Hsieh, *J. Biol. Chem.*, 1997, **272**, 23811.
11 B. Ames, F. D. Lee and W. E. Durston, *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 782.