DOI:
10.1016/j.bmcl.2016.10.022

Document Version
Peer reviewed version

DOI: http://dx.doi.org/10.1016/j.bmcl.2016.10.022
Reference: BMCL 24325

To appear in: *Bioorganic & Medicinal Chemistry Letters*

Received Date: 17 June 2016
Revised Date: 6 October 2016
Accepted Date: 9 October 2016

Please cite this article as: M. Jackson, P.J., Rahman, K.M., Thurston, D.E., The use of molecular dynamics simulations to evaluate the DNA sequence-selectivity of G-A cross-linking PBD-duocarmycin dimers, *Bioorganic & Medicinal Chemistry Letters* (2016), doi: http://dx.doi.org/10.1016/j.bmcl.2016.10.022

Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.

The use of molecular dynamics simulations to evaluate the DNA sequence-selectivity of G-A cross-linking PBD-duocarmycin dimers

Leave this area blank for abstract info.

Paul J. M. Jackson [a, b], Khondaker M. Rahman [a, b] and David E. Thurston [a, b]

[a] *Femtogenix Ltd, Britannia House, 7 Trinity Street, London SE1 1DB, United Kingdom.*

[b] *Institute of Pharmaceutical Science, Faculty of Life Sciences & Medicine, King's College London, Britannia House, 7 Trinity Street, London SE1 1 DB, United Kingdom.*

Bioorganic & Medicinal Chemistry Letters
journal homepage: www.elsevier.com

The use of molecular dynamics simulations to evaluate the DNA sequence-selectivity of G-A cross-linking PBD-duocarmycin dimers

Paul J. M. Jackson [a, b], Khondaker M. Rahman [a, b] and David E. Thurston [a, b]
[a] *Femtogenix Ltd, Britannia House, 7 Trinity Street, London SE1 1DB, United Kingdom*
[b] *Institute of Pharmaceutical Science, King's College London, Britannia House, 7 Trinity Street, London SE1 1DB, United Kingdom*

ARTICLE INFO

Article history:
Received
Revised
Accepted
Available online

Keywords:
DNA cross-linking
Pyrrolobenzodiazepine
Duocarmycin
PBD-CBI dimer
Molecular dynamics

ABSTRACT

The pyrrolobenzodiazepine (PBD) and duocarmycin families are DNA-interactive agents that covalently bond to guanine (G) and adenine (A) bases, respectively, and that have been joined together to create synthetic dimers capable of cross-linking G-G, A-A, and G-A bases. Three G-A alkylating dimers have been published to date, with defined DNA-binding sites proposed for two of them. In this study we have used molecular dynamics simulations to elucidate preferred DNA-binding sites for the three published molecular types. For the PBD-CPI dimer UTA-6026 (**1**), our simulations correctly predicted its favoured binding site (*i.e.*, 5'-C(G)AATTA-3') as identified by DNA cleavage studies. However, for the PBD-CI molecule ("Compound 11", **3**), we were unable to reconcile the results of our simulations with the reported preferred cross-linking sequence (5'-ATTTTCC(G)-3'). We found that the molecule is too short to span the five base pairs between the A and G bases as claimed, but should target instead a sequence such as 5'-ATTTC(G)-3' with two less base pairs between the reacting G and A residues. Our simulation results for this hybrid dimer are also in accord with the very low interstrand cross-linking and *in vitro* cytotoxicity activity reported for Compound 11. Although a preferred cross-linking sequence was not reported for the third hybrid dimer ("27eS", **2**), our simulations predict that it should span three base pairs between covalently reacting G and A bases (*i.e.*, 5'-GTAT(A)-3'.

Despite their age, DNA cross-linking agents such as cisplatin[1] and bendamustine[2] remain useful agents in the clinic to treat a number of different cancer types[3]. Through the years, many novel cross-linking agents have been designed, synthesized and evaluated[4]'[5] in an attempt to (a) improve on cross-linking efficiency[6], (b) modify the sequence-selectivity of cross-linking[7], (c) change the distribution of inter- *versus* intrastrand cross-links [8-10], and/or (d) target different base pair combinations[11-13]. For example, the naturally occurring pyrrolobenzodiazepine (PBD) molecules, derived from *Streptomyces* species[14-18], are characterized by an N10-C11 imine group that enables formation of a covalent bond to the C2-amino group of a guanine base[19], and these have been successfully incorporated into PBD dimer structures such as DSB-120[20] and SJG-136[21] **(Figure 1A)** that can form sequence-selective G-G cross-links in the DNA minor groove. The cytotoxicity of these molecules is thought to be directly related to the formation of interstrand cross-links (in this instance at Pu-GATC-Py sites[22, 23]), and members of the PBD dimer class are now being used as payloads for antibody-drug conjugates (ADCs)[24, 25]. Evidence is accumulating to suggest additional mechanisms of action including the formation of intrastrand cross-links[9, 10, 26] and mono-adducts, the inhibition of transcription factor binding[27, 28] and the inhibition of enzymes such as endonucleases [29, 30] and RNA polymerase[31].

A related but similar example is the duocarmycin family of DNA-interactive agents (*e.g.*, CC1065), also isolated from *Streptomyces* species[32-34] and used as payloads for ADCs[35]. Members of this family covalently bond to the N3-position of adenine bases in the DNA minor groove[36]. In a similar manner to other DNA minor-groove binding molecules such as the polyamides[37], the ability of these molecules to initially bind to DNA is thought to relate to their curved shape which allows them to embed in AT tracts of the minor groove where they initially form stabilizing non-covalent interactions (*e.g.*, van der Waals interactions) with the minor groove walls[38, 39]. However, unlike the PBDs, duocarmycin-type molecules are thought to be "activated" through a DNA-binding-induced conformational change whereby the vinylogous amidic conjugation is disrupted, bringing the cyclopropane unit into alignment with the adjacent cyclohexadienone system, thus activating it for nucleophilic attack, a process known as shape-dependent catalysis[38-41]. Thus, members of the duocarmycin family are thought to be chemically unreactive until they reach their DNA target and undergo this activation process[42]. These molecules (*i.e.*, as their component CPI units) have also been joined through a chemical linker to create analogues such as Bizelesin[43]

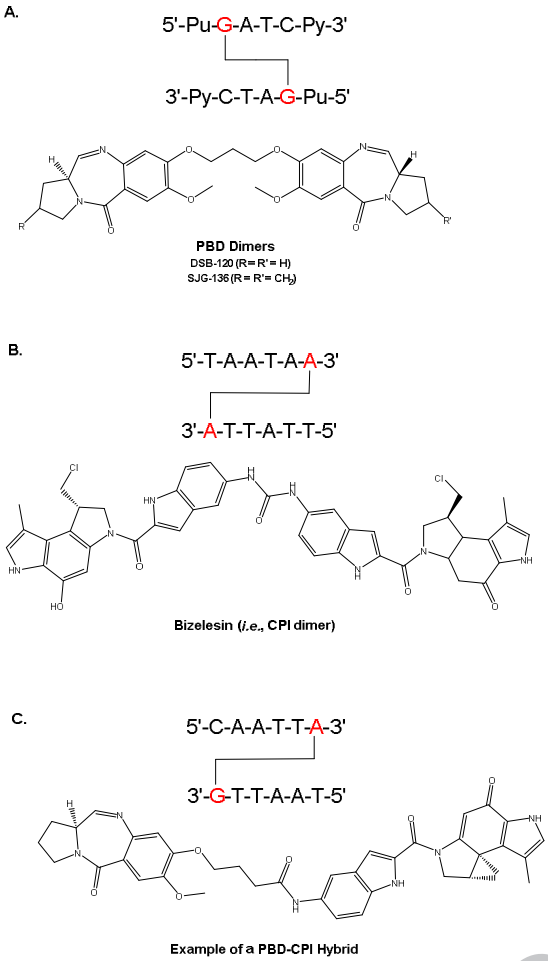
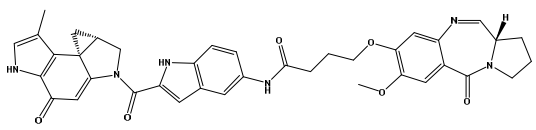


Figure 1. **A**. The PBD Dimers (*e.g.*, DSB-120 and SJG-136) which can form DNA minor-groove interstrand G-G cross-links; **B**. A CPI Dimer (*e.g.*, Bizelesin) which can form minor-groove interstrand A-A cross-links; **C**. An example of a CPI-PBD hybrid molecule that can cross-link A-G base pairs. The mode of interstrand cross-linking and likely base-pair span is shown above each family structure (covalently modified bases in red).

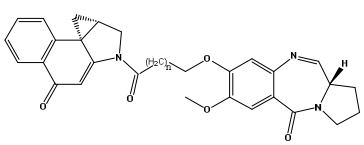
(**Figure 1B**) which is capable of forming sequence-selective A-A interstrand cross-links in the DNA minor groove. Other CPI dimers have been produced[44], joined through their N3-positions.

The incorporation of both GC-interactive and AT-interactive moieties into symmetric G-G and A-A dimers led to the concept of joining PBD and CPI units together to create hybrid molecules capable of forming cross-links to both G and A bases simultaneously (**Figure 1C**), which are the subject of this study. Three types of these molecules have been reported in the literature[45][46][47], and their structures are shown in **Figure 2**.

The objective of this study was to use, for the first time, molecular dynamics simulations to predict the preferred DNA recognition sequences for the three published types of G-A cross-linking agents[45][46][47] shown in **Figure 2**, and then compare these predictions with the cross-linking sites proposed by the authors. Hurley and co-workers published the first example of a CPI-PBD dimer (UTA-6026, **1, Figure 2**) in 2001[45], reporting that it cross-links the sequence 5'-C(G)AATTA-3' (covalently-modified bases underlined; brackets denote base on opposite strand), has a CL_{50} value (a measure of cross-linking potency) of between 1-10μM in calf thymus DNA, and an IC_{50} as low as 0.047 nM in some tumour cell lines (*i.e.*, colon SW480). Interestingly, in their gel studies, cross-linking was not complete even at 100μM, suggesting that the molecule is a relatively inefficient cross-linking agent. In 2003, Denny and co-workers reported[46] an extensive study of PBD-CPI hybrids which differed from the Hurley dimer in containing simple methylene linkers between the CPI and PBD units, and these molecules are now being developed as payloads for antibody-drug conjugates through Genentech. Their most potent molecule *27eS* (**2**, **Figure 2**) had a CL_{50} value of ~0.01μM toward pcDNA3 plasmid DNA, and an IC_{50} as low as 0.0078 nM in some tumour cell lines (*i.e.*, EMT6). The molecule also had limited antitumour activity in human tumour xenograft mouse models. The authors provided some evidence that G-A cross-linking was occurring, but did not suggest a specific sequence to which **2** should bind. Finally, in 2006 Lee and co-workers synthesized the *seco*-amino-CI-PBD hybrid 'Compound 11' (**3, Figure 2**) which they reported to provide a very low degree of cross-linking toward pBR-322 DNA at 10 μM, and to have an IC_{50} of 0.56 μM in P815 cells after 72 hours exposure. Based on the results of a thermally-induced cleavage assay, they reported that **3** forms cross-links at the 5'-AATTTTCC(G)-3' sequence.



1 (UTA-6026)



2 (27eS), n=4

Figure 2. Structures of the three published families of PBD-(CPI/seco-CI) hybrid molecules capable of interstrand cross-linking between G and A base pairs in the DNA minor groove: The Hurley (**1**)[45], Denny (**2**)[46] and Lee (**3**)[47] hybrid dimers.

During the past six years we have developed molecular dynamics simulation methodologies[11, 13, 48] that allow us to accurately predict and interpret how both mono-alkylating and DNA cross-linking agents of this type interact in the minor groove of DNA. As the hybrid dimer **3** reported by Lee [47] had a significantly lower DNA cross-linking potency and *in vitro* cytotoxicity compared to the Hurley (**1**) and Denny (**2**) dimers, we decided to use our simulation methodologies to compare the interaction of the three molecules at their proposed DNA recognition sites. We report here that, although the DNA-interactivity of hybrid dimers **1** and **2** correlate well with their cross-linking activity and cytotoxicity, the Lee dimer **3** does not appear to have the correct 3-dimensional shape to fit into the DNA minor groove, and, in addition, its two alkylating moieties (*i.e.*, the PBD and *seco*-amino-CI units) cannot orient into the appropriate alignment for covalent bonding with the relevant DNA bases on either the opposite or same strands of DNA. More significantly, the MD simulations predict that the linker is of insufficient length to allow the PBD and seco-amino-CI units to reach the G and A DNA bases that the authors claim to be covalently modified.

MD simulations were undertaken over a 10ns duration (extended methodology presented in **Supplementary Material**). When comparing the three molecules, RMSD calculations between the energy-minimized structures of each adduct and the lowest energy snapshot derived from MD simulations provided a reflection of the DNA disorder observed during simulations. Interestingly, the RMSD calculation results correctly predicted the experimentally-derived binding site for **1**, suggesting that RMSD values may be useful as indicators of DNA-affinity and cross-link span.

The CPI-PBD hybrid UTA-6026 (**1**) was designed by Hurley and co-workers using a molecular modeling approach (details not published) to span six base-pairs, covalently binding to a guanine (through the PBD) and an adenine (through the CPI moiety)[45]. An indole group was included as part of the central linking moiety to increase interactivity with functional groups in the minor groove through van der Waals interactions. Hurley and co-workers predicted that, for their chosen sequence, the optimal methylene linker length between the C8-position of the PBD and the amide functionality (connecting to the indole ring) should be n = 3 to allow the PBD and CPI units to lie in the correct orientation to alkylate DNA co-operatively[45]. Non-denaturing gel electrophoresis was used to investigate the cross-linking ability of **1**, and a thermal cleavage gel-based assay was performed to identify adenine cleavage points. Together, these data indicated cross-link formation at two different 5'-C(G)AATTA-3' sites (alkylated bases underlined).

In the present study, molecular dynamics simulations were carried out over a 10ns time-scale to confirm the proposed six base-pair span of **1**, and the interstrand cross-link at 5'-GCC(G)AATTAGC-3' as derived from the published experimental data. The simulation results were in full agreement with the experimental data. When mono-covalently bound through the CPI moiety alone to position A8 of this sequence, the PBD was located close enough to G18 for nucleophilic attack and covalent attachment to occur. Furthermore, when both the PBD and CPI units were covalently bound to G18 and A8, respectively, in the same sequence, negligible distortion of the DNA was observed (*i.e.*, RMSD = 0.78Å, **Table 1**), with the molecule well accommodated in the minor groove (**Figure 3**).

Simulations with **1** covalently bound to sequences one base shorter and one or two bases longer (*i.e.*, 5'-CAATA-3', 5'-CAAATTA-3' and 5'-CAAAATTA-3') gave higher RMSD values of 1.96, 1.00 and 3.04, respectively, thus supporting the preferred 5'-CAATA-3' sequence (**Table 1**). It is noteworthy that the difference in RMSD values between the preferred and shorter sequence (*i.e.*, 0.78 *vs* 1.96) was much greater than between the preferred and one base-pair longer sequences (*i.e.*, 0.78 *vs* 1.00), suggesting that if the preferred sequence were to be extended by one base-pair, the ligand may still be able to form adducts. It is also interesting that the differences between all of these RMSD values for **1** (*i.e.,* 0.78, 1.96, 1.00 and 3.04Å) are far smaller than those observed for **2** (*i.e.,* 0.93, 2.42 and 6.45Å). **2** contains a methylene chain between alkylating moieties, whereas **1** contains an indole group, which may stabilize the adduct to a greater degree than the flexible methylene linker.

Non-covalently bound simulations provided a similar pattern to dimers **2** and **3** in that the molecule appeared to be restrained over a guanine for the duration of the simulation, but clearly showed that six base-pairs were being spanned **(Figure 3).** Also, because this sequence span is independent of the length of the DNA sequences being investigated, the Free Energy of Binding calculations are similar in each case, and do not reflect the more meaningful observations based on RMSD calculations from covalently bound simulations.

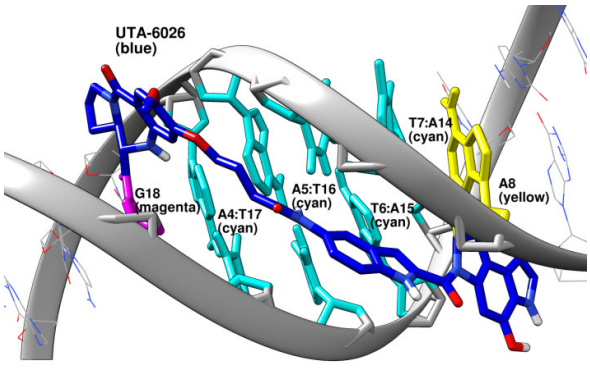


Figure 3. Snapshot of an MD simulation of the Hurley hybrid PBD-CPI dimer (**1**) (blue) with its PBD and CPI units covalently bound to the G18 (magenta) and A8 (yellow) bases, respectively, of the sequence 5'-GCC**(G)AATTA**GC-3', with excellent accommodation in the minor groove and little distortion of the central AT base-pairing (cyan).

Based on a thermally-induced DNA cleavage assay of **3**, Lee and co-workers[47] proposed that it has a seven base-pair cross-linking recognition site of 5'-AATTTTCC(G)-3', and also reported a number of mono-alkylation events at 5'-TAA-3' sites

within linearized pBR322 plasmid DNA. Based on this, we initially conducted an MD simulation of the published adduct (*i.e.* **3** covalently cross-linked to the seven base-pair duplex sequence 5'-GAATTTTCC(G)C-3'), with the introduction of a 'GC-lock' (*i.e.,* terminal guanine bases) to ensure that the DNA duplex remained intact throughout the simulation. As CPI/CBI-type molecules are known to be more reactive toward DNA than PBD molecules[46], the simulations involved initial covalent attachment of just the *seco*-amino-CI unit, leaving the PBD unreacted. This allowed us to investigate the span of the entire hybrid molecule. Next, the DNA sequence was gradually reduced in length to investigate the shortest possible adduct which was determined to be 5'-ATTC(G)-3' (**Table 1**). An identical set of simulations was then undertaken with both PBD and CI covalently bound to guanine and adenine residues, respectively, within the same set of sequences. Finally, simulations were undertaken with **3** bound non-covalently in the DNA minor groove to enable calculation of free energy of binding of the ligand to DNA **(Table 1)**.

It was immediately evident from the simulations that **3** was unable to span the seven base-pair cross-link proposed by Lee and co-workers (*i.e.,* 5'-ATTTTCC(G)-3'). In particular, simulations with the molecule *bis*-covalently bonded to this sequence indicated that severe distortion of the DNA helix would occur with disruption of base-pairing and extensive minor groove widening (**Figure 4A** and **4C**). This result was also supported by RMSD plots of the adduct over the duration of the simulation (**Supplementary Material**), and RMSD calculations between the lowest energy conformation of the molecule derived from MD simulations and the lowest energy conformation after energy minimization (pre-MD simulation). RMSD values were calculated to be 2.45Å, suggesting a high degree of distortion of the ligand:DNA complex. Crucially, simulations with only the

Table 1. Free Energy of Binding and RMSD calculations for interaction of the PBD-duocarmycin dimer hybrids **1**-**3** with the DNA sequences shown, as calculated over the duration of the MD simulations.

Hybrid Dimer	**Molecule Span (number of base-pairs)**	**DNA Sequence** (Sequence being analysed is highlighted in bold and red)	**Free Energy of Binding Calculations** (generated from non-covalently bound simulations) (kcal/mol)	**RMSD Calculations** (generated from covalently bound simulations) (Å)	**Base-pairing maintained during simulation?** (Yes/No)
1	5-6	5'-GC**CAATA**GC-3'	-55.45	1.96	Y
		5'-GC**CAATTA**GC-3'	-54.15	0.78	Y
		5'-GC**CAAATTA**GC-3'	-54.40	1.00	Y
		5'-GC**CAAAATTA**GC-3'	-54.32	3.04	N
2	4-5	5'-GC**GTAT**GC-3'	-41.11	0.93	Y
		5'-GC**GTAAT**GC-3'	-41.97	2.42	N
		5'-GC**GTAAAT**GC-3'	-38.93	6.45	N
		5'-GC**GTAAAAT**GC-3'	-41.60	5.23	N
3	4-5	5'-GA**ATTC**G-3'	-48.78	1.82	N
		5'-GA**ATTTC**G-3'	-51.41	1.15	N
		5'- GA**ATTTTC**G-3'	-52.84	5.72	N
		5'-GA**ATTTTCC**G-3'	-52.63	8.45	N

seco-amino-CI unit covalently bonded to the adenine indicated that the PBD moiety is then situated directly over the T6:A15 base-pair **(Figure 4B** and **4C)**, at least two base-pairs away from the reacting guanine proposed by Lee and co-workers[49], thus rendering the proposed second alkylation event unlikely.

A similar pattern was evident in simulations involving the six base-pair sequence 5'-ATTTTC(G)-3', where mono-covalent simulations indicated difficulty in cross-link formation, and dual-covalently bound simulations suggested significant DNA distortion (indicated by a RMSD value of 5.72 (**Table 1**). Although mono-covalently bound simulations suggested a potential cross-link at the four base-pair sequence 5'-ATTC(G)-3', dual-covalently bound simulations (RMSD = 1.82) suggested that the five base-pair sequence 5'-ATTTC(G)-3' should form the most favoured adduct. This was supported by the RMSD calculations (**Table 1)**, which indicated less distortion for both the four and five base-pair adducts (*i.e.*, RMSD = 1.82 and 1.15, respectively), suggesting that both might form.

Overall these simulations highlighted the importance of DNA breathing in the accommodation of ligands in the minor groove. Our observation that **3** is likely to cross-link the five base-pair DNA sequence 5'-ATTTC(G)-3' rather than the extended sequence 5'-ATTTTCC(G)-3' as suggested by the authors was further supported by the proximal location of the unbound PBD imine close to the reacting guanine in the mono-alkylated adduct in the former sequence. Although significant distortion was observed in simulations when dual-covalently bound to the idealized DNA sequence, models suggested that the molecule is likely to span five base-pairs (see **Supplementary Material** for molecular model). Furthermore, as 5'-ATTTC(G)-3' is not obvious as a potential binding site from the published gel studies, it is likely that the authors were observing mono-alkylation rather than cross-linking events in their experiments.

Non-covalently bound simulations of **3** were also not reflective of the reported experimental results. Previous studies have shown that adenine-binding molecules such as the CPI family direct the sequence targeting of unsymmetrical PBD-CPI hybrids[46]. However, in this case, the non-covalently bound simulations suggested that the PBD moiety directs binding due to the formation of a sequence-selective H-bond between the lone pair of the N10 of the PBD and the exocyclic amine of a guanine residue. As a result, in the case of every non-covalently bound simulation undertaken, and thus every Free Energy of Binding calculation, the PBD moiety remained restrained over the intended reacting guanine (*i.e.*, G14 on the reverse strand in the case of 5'-GAATTTTC(**G**)C-3'), spanning DNA in an A-ring-5' orientation in the process. As **3** spans four base-pairs, addition of an extra thymine base to the centre of the binding site does not alter the sequence with which the molecule is interacting, as the PBD is restrained over a G:C base-pair. As a result, calculated free energy of binding values are similar for all sequences analysed.

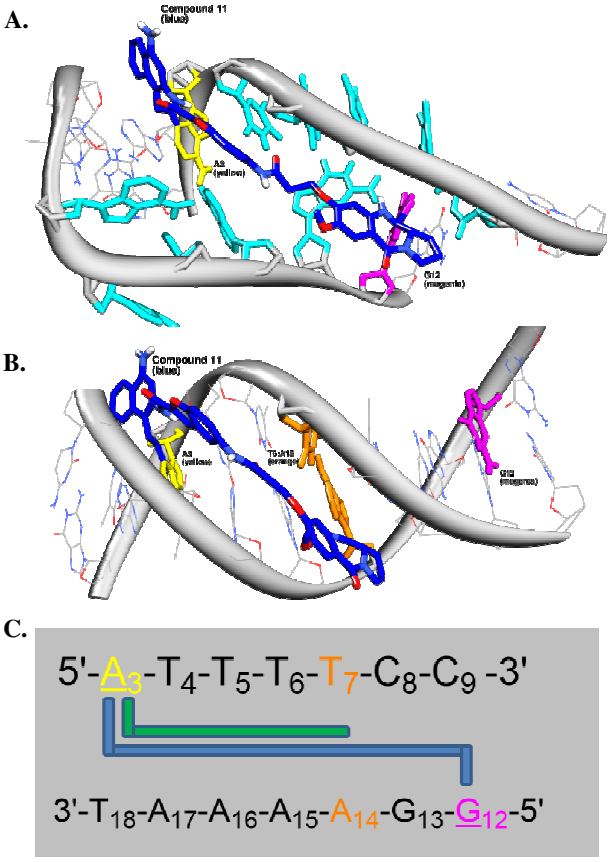


Figure 4. **A:** The PBD-CI hybrid dimer **3** of Lee and co-workers[47] covalently bound to an adenine (yellow) and guanine (magenta) (underlined) through the CI and PBD units, respectively, in the sequence 5'-GCATTTTCC(G)C-3' proposed by the author, illustrating extensive base-pair distortion (cyan) and minor groove widening; **B:** **3** mono-covalently bound to A3 (yellow, underlined) of 5'-GAATTTTCCC-3' through its *seco*-amino-CI moiety. The molecule spans four to five base-pairs (green rod in schematic **C**), with the PBD moiety ideally placed for covalent attack on either of the bases at the T7:A14 positions (orange) if they were mutated to guanines. The G12 (magenta), suggested by Lee and co-workers to be the target base for covalent interaction of the PBD component of the molecule (blue rod in schematic **C**) would lead to significant distortion of the DNA helix (see **A**).

Bioinformatics analysis of the linearized DNA plasmid pBR322 suggests 20 potential binding sites for **3** (from within a total of 4361 bases) based on the proposed preferred five base-pair 5'-AWWWC(G)-3' binding site (where W represents adenine or thymine, G represents guanine and C is cytosine) from our modeling studies. The relative prevalence of the sequence 5'-AWWWC(G)-3' suggests that a significant band in the cross-linking gel relating to this sequence should have been observed. However, as this was not the case, it is likely that other factors may contribute to the lack of interaction of **3** with DNA. In this context, it is noteworthy that, unlike the potent PBD dimers in the literature (**Figure 1A**)[50], **3** possesses a linker between the C8-position of the PBD and the amidic carbonyl based on two methylene groups. Symmetrical PBD dimers in which the linker is odd-numbered (*i.e.* n = 3 or 5) are known to be more DNA interactive and cytotoxic (*e.g.*, 10^3 more potent in some cancer cell lines) than even-numbered equivalent molecules (*i.e.* n = 2 or 4), as the alkylating moieties are more-appropriately oriented for the alkylation events[51]. Therefore, we considered whether this effect could have contributed to the poor cross-linking ability observed for **3** by carrying out simulations on analogues containing an alternative trimethylene linker (n=3) (data not shown). This led us to conclude that these rules derived from the PBD dimers may also apply to asymmetric hybrid dimer molecules, whereby odd-numbered methylene linkers appear to be preferential for *bis*-alkylation events to occur.

For the PBD-CPI hybrids produced by Tercel and co-workers[46], their thermal cleavage experiments suggested that a number of mono-alkylation events occurred between the CPI moiety of **2** and adenine bases, and results from their Comet assay analysis suggested greater interstrand cross-linking potency for some of their hybrid dimers compared to the PBD dimer DSB-120[46]. However, potential DNA cross-linking sites were not identified in their study. Therefore, we used our molecular dynamics simulations to try to identify potential cross-linking sites.

Starting with a four base-pair sequence (5'-GTAT-3'), an increasing number of adenine-thymine base-pairs were inserted between the 5'-T and 5'-A bases (*i.e.*, 5'-GT(A)$_n$T-3'), increasing the total span of base-pairs up to a maximum of seven (*i.e.*, 5'-GTAAAAT-3'). A:T base-pairs were selected as the central components in an effort to maximise non-covalent interaction with functional groups in the DNA minor groove, and minimize potential steric hindrance from the C2-amino groups of guanine bases. In a similar manner to the simulations undertaken on the other hybrid dimers, the interaction of **2** with DNA was simulated in three ways; non-covalently bound, mono-alkylated with the CPI moiety covalently bound to an adenine base, and dual-alkylated. These simulations suggested that the four base-pair sequence 5'-GTAT-3' is the most favoured cross-linking site (**Figure 5**). This was mainly reflected in the *bis*-covalently bound simulations, where little distortion of the DNA structure was observed (*i.e.*, RMSD = 0.93Å, **Table 1**), and base-pairing was maintained. Introduction of additional base pairs beyond the established four base-pair optimum resulted in increased DNA disorder, which was reflected in higher RMSD values (*e.g.*, RMSD values of up to 6.45Å in the longer sequences analyzed, **Table 1**).

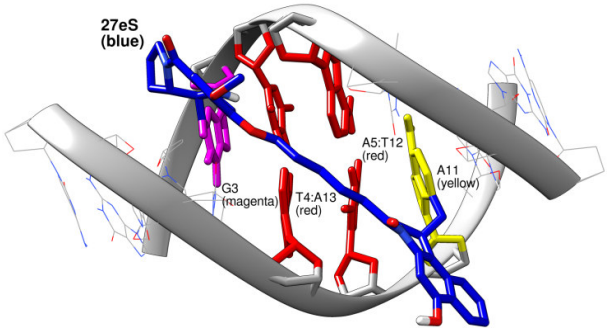


Figure 5. The PBD-CPI hybrid dimer **2** ("27eS") synthesised by Denny and co-workers[47] (blue) with the CPI and PBD moieties covalently bound to both an adenine (yellow) and guanine (magenta), respectively, in the idealized sequence 5'-GC**G**T**A**T(**A**)GC-3' predicted through molecular dynamics simulations, illustrating the excellent accommodation of the molecule in the minor groove for this sequence.

These results showed that 5'-GWWT-3' (where W represents A or T) should be the most favoured cross-linking site for **2.** In a similar manner to simulations of **3,** Free Energy of Binding

calculations did not reflect the predicted cross-linking sites. During simulations between **2** and each sequence, a sequence-selective hydrogen bond was formed between the C2-amino group of guanine and the lone pair of the N10 of the PBD, thereby restraining **2** over a guanine base for the duration of the simulation. Therefore, as the span of the molecule is four to five base-pairs, there was little difference between each sequence from the simulation perspective (**Table 1**), and so negligible differences in Free Energy of Binding values were observed.

Finally, it is known that symmetrical PBD dimers in which the linker attached to the C8/C8'-positions is odd-numbered in length (*e.g.*, n = 3 or n = 5) have a higher DNA binding affinity and are more cytotoxic (*e.g.*, 10^3 more potent in some cancer cell lines) than even-numbered equivalent molecules (*e.g.*, n = 2 or 4). This is thought to be due to the alkylating moieties being in a more optimal orientation and alignment for alkylation events to occur in the case of odd-numbered linkers[51]. The hybrid dimers synthesized by Tercel and co-workers[46] appear to follow this rule. For example, they synthesized and evaluated a series based on CPI and PBD moieties separated by a methylene linker (*i.e.*, based on **2**, **Figure 2**), with differentiation achieved through variation of linker length (*i.e.*, n = 1-5). Compounds with odd-numbered linker lengths (in this case n = 3 and n = 5) significantly out-performed those with even-numbered linker lengths, providing a cross-linking hierarchy of $n = 2 < 1 < 4 < 5 \leq 3$. This experimental observation is in accord with our models of **3,** which suggest that the analogue with a linker of n = 3 should be more DNA-interactive than the n = 2 analogue.

It is important to acknowledge that this molecular dynamics simulation approach may have a number of limitations. For example, one problem is that the covalent binding energy of forming just one covalent bond between any ligand and DNA is so large (*i.e.*, ~75 kcal/mol per C-N bond) that there is abundant energy available to radically distort the DNA structure, and for cross-linking agents which form two covalent bonds even more free energy is available (*i.e.*, ~150 kcal/mol). This covalent binding free energy is not encompassed in the molecular dynamics simulation paradigm, and so this approach may not be able to predict the formation of highly distorted DNA adducts (*e.g.*, those produced by cross-linking agents such as *cis*-platin)[3]. One practical problem when carrying out a molecular dynamics simulation is the need to establish which of the many conformations formed during the simulations are potentially relevant in real life. In the case of ligand-DNA interactions, if the simulations produce solutions that strongly resemble canonical B-form DNA with the ligand fitting perfectly in the minor groove and making good Van der Waals interactions in all directions, then it might be reasonable to assume that such "equilibrium" configurations represent plausible structures in real life. Although we cannot be completely confident that the conformation chosen post-minimization (*i.e.*, before undertaking simulation) is the most biologically relevant, it is evident from our results that measuring the degree of distortion over the duration of a simulation can provide an insight into cross-linking efficiency. In the case of the molecules studied here, cross-linking potency, cytotoxicity and degree of distortion of ligand:DNA adducts are experimentally correlated, and it is gratifying that our MD simulations appear to follow this trend. Interestingly, in the case of the PBD dimers (**Figure 1A**), their significant cytotoxicity in cells is thought to be due to their perfect fit within the minor groove and consequent lack of distortion upon covalent cross-linking, thus avoiding the attention of DNA repair enzymes[17].

In summary, we report the first molecular dynamics simulations of the interaction of PBD-CPI hybrid dimers with DNA. In the case of the Hurley hybrid dimer UTA-6026 (**1**)[45], our simulations correctly predicted the cross-link shown to form experimentally in a DNA cleavage assay, and this served to validate our modeling protocols for use with the other G-A cross-linking agents, **2** and **3**. However, our modeling studies illustrated that the cross-link proposed by Purnell and co-workers for **2** is unlikely to occur, and that the results of their reported gel studies most likely relate to mono-alkylation rather than cross-linking events. Also, our calculations demonstrate that RMSD values between the lowest energy derived from energy minimization studies and those obtained through molecular dynamics simulations are a valuable tool in predicting cross-link formation, and may be useful in the design of more potent next-generation PBD-(CBI/CPI/CI) dimers targeted to specific DNA sequences.

Acknowledgement

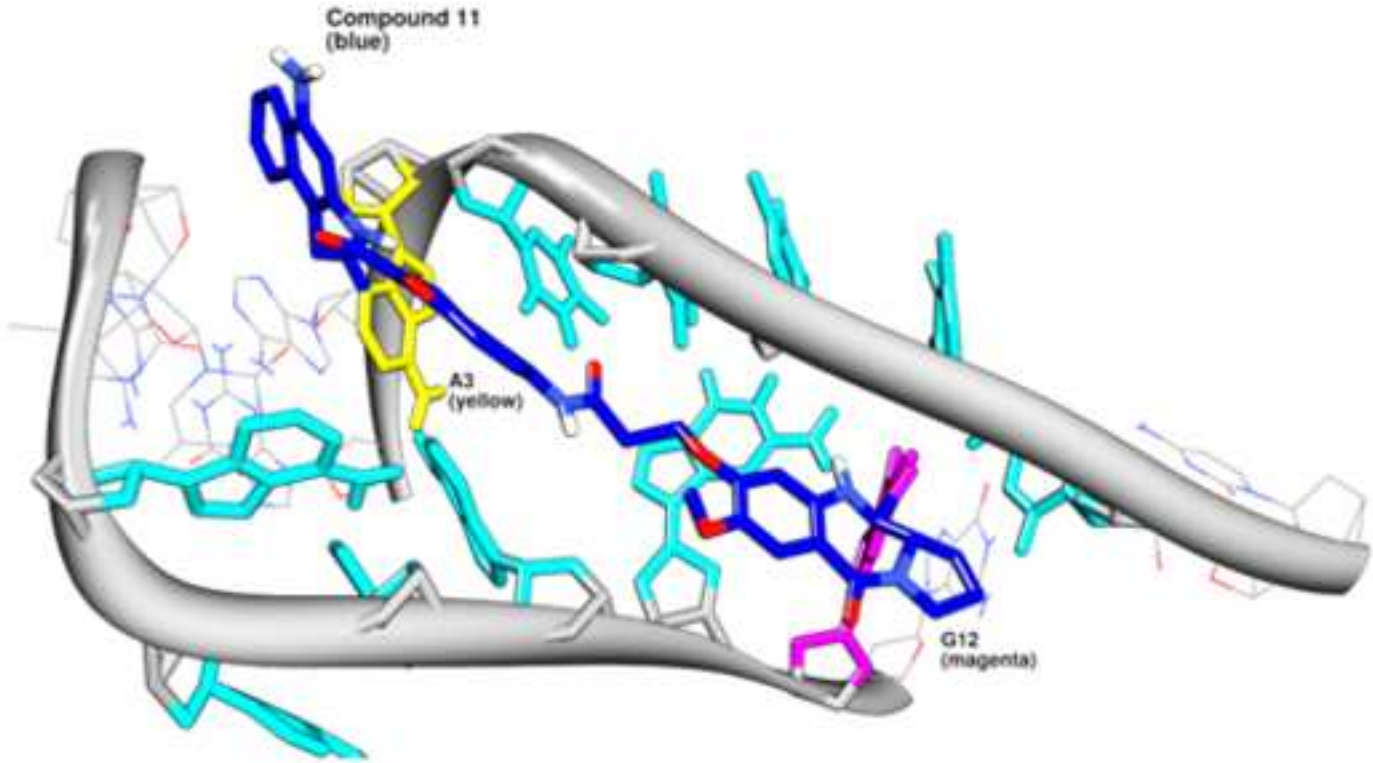
The Frederick Gardener Cottrell Foundation (Arizona, USA) is thanked for financial support of this work.

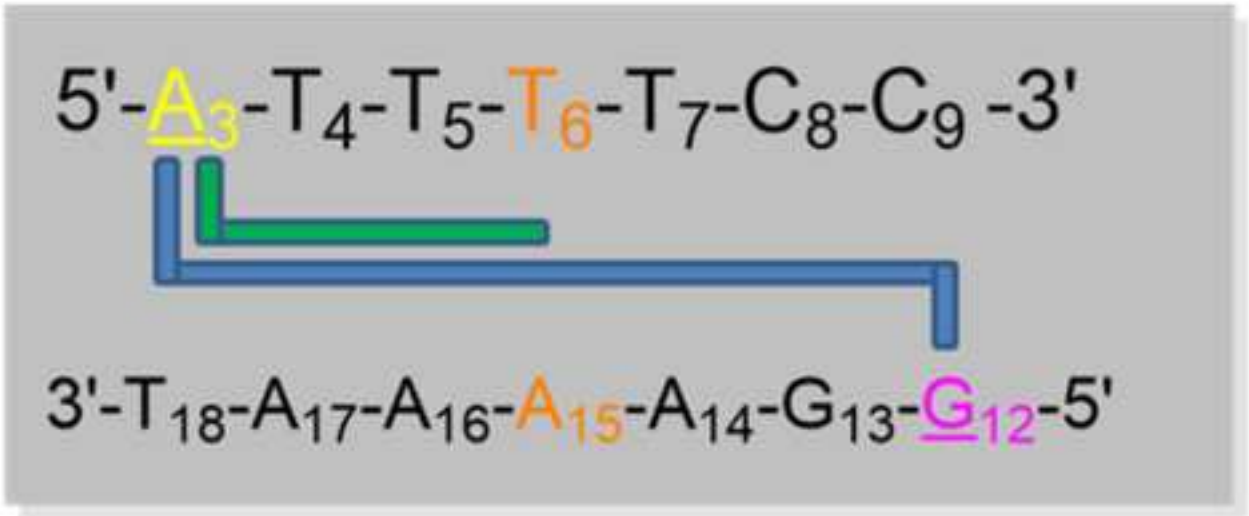
References and Notes

[1] B. Rosenberg, L. Van Camp, T. Krigas, Inhibition of Cell Division in Escherichia coli by Electrolysis Products from a Platinum Electrode, Nature, 205 (1965) 698-699.
[2] A. Gilman, The initial clinical trial of nitrogen mustard, Am J Surg, 105 (1963) 574-578.
[3] D.E. Thurston, Chemistry and Pharmacology of Anticancer Drugs, CRC Press (Taylor & Francis), Boca Raton, Florida, USA, 2006.
[4] L. Brulikova, J. Hlavac, P. Hradil, DNA interstrand cross-linking agents and their chemotherapeutic potential, Current medicinal chemistry, 19 (2012) 364-385.
[5] Y. Huang, L. Li, DNA crosslinking damage and cancer - a tale of friend and foe, Translational cancer research, 2 (2013) 144-154.
[6] S.T. Sigurdsson, S.M. Rink, P.B. Hopkins, Affinity crosslinking of duplex DNA by a pyrrole-oligopeptide conjugate, J. Am. Chem. Soc., 115 (1993) 12633-12634.
[7] H. Rosado, K.M. Rahman, E.-A. Feuerbaum, J. Hinds, D.E. Thurston, P.W. Taylor, The minor groove-binding agent ELB-21 forms multiple interstrand and intrastrand covalent cross-links with duplex DNA and displays potent bactericidal activity against methicillin-resistant Staphylococcus aureus, Journal of Antimicrobial Chemotherapy, 66 (2011) 985-996.
[8] J. Malina, O. Novakova, M. Vojtiskova, G. Natile, V. Brabec, Conformation of DNA GG Intrastrand Cross-Link of Antitumor Oxaliplatin and Its Enantiomeric Analog, Biophysical Journal, 93 (2007) 3950-3962.
[9] K.M. Rahman, A.S. Thompson, C.H. James, M. Narayanaswamy, D.E. Thurston, The Pyrrolobenzodiazepine Dimer SJG-136 Forms Sequence-Dependent Intrastrand DNA Cross-Links and Monoalkylated Adducts in Addition to Interstrand Cross-Links, Journal of the American Chemical Society, 131 (2009) 13756-13766.
[10] K.M. Rahman, C.H. James, D.E. Thurston, Effect of Base Sequence on the DNA Cross-Linking Properties of Pyrrolobenzodiazepine (PBD) Dimers, Nucleic Acids Research, 39 (2011) 5800-5812.
[11] F. Brucoli, R.M. Hawkins, C.H. James, P.J. Jackson, G. Wells, T.C. Jenkins, T. Ellis, M. Kotecha, D. Hochhauser, J.A. Hartley, P.W. Howard, D.E. Thurston, An Extended Pyrrolobenzodiazepine-Polyamide Conjugate with Selectivity for a DNA Sequence Containing the ICB2 Transcription Factor Binding Site, Journal of medicinal chemistry, 56 (2013) 6339-6351.
[12] M. Kotecha, J. Kluza, G. Wells, C.C. O'Hare, C. Forni, R. Mantovani, P.W. Howard, P. Morris, D.E. Thurston, J.A. Hartley, D. Hochhauser, Inhibition of DNA Binding of the NF-Y Transcription

Factor by the Pyrrolobenzodiazepine-Polyamide Conjugate GWL-78, Mol. Cancer Ther., 7 (2008) 1319-1328.
[13] K.M. Rahman, P.J.M. Jackson, C.H. James, B.P. Basu, J.A. Hartley, M. de la Fuente, A. Schatzlein, M. Robson, R.B. Pedley, C. Pepper, K.R. Fox, P.W. Howard, D.E. Thurston, GC-Targeted C8-Linked Pyrrolobenzodiazepine–Biaryl Conjugates with Femtomolar in Vitro Cytotoxicity and in Vivo Antitumor Activity in Mouse Models, Journal of medicinal chemistry, 56 (2013) 2911-2935.
[14] D. Antonow, D.E. Thurston, Synthesis of DNA-interactive pyrrolo[2,1-c][1,4]benzodiazepines (PBDs), Chem Rev, 111 (2011) 2815-2864.
[15] L. Cipolla, A.C. Araujo, C. Airoldi, D. Bini, Pyrrolo[2,1-c][1,4]benzodiazepine as a scaffold for the design and synthesis of anti-tumour drugs, Anticancer Agents Med Chem, 9 (2009) 1-31.
[16] B. Gerratana, Biosynthesis, synthesis, and biological activities of pyrrolobenzodiazepines, Med Res Rev, 32 (2012) 254-293.
[17] J.A. Hartley, The development of pyrrolobenzodiazepines as antitumour agents, Expert Opin Investig Drugs, 20 (2011) 733-744.
[18] A. Kamal, K.L. Reddy, V. Devaiah, N. Shankaraiah, D.R. Reddy, Recent advances in the solid-phase combinatorial synthetic strategies for the benzodiazepine based privileged structures, Mini Rev Med Chem, 6 (2006) 53-69.
[19] L.H. Hurley, T. Reck, D.E. Thurston, D.R. Langley, K.G. Holden, R.P. Hertzberg, J.R. Hoover, G. Gallagher, Jr., L.F. Faucette, S.M. Mong, et al., Pyrrolo[1,4]benzodiazepine antitumor antibiotics: relationship of DNA alkylation and sequence specificity to the biological activity of natural and synthetic compounds, Chem Res Toxicol, 1 (1988) 258-268.
[20] D.S. Bose, A.S. Thompson, J.S. Ching, J.A. Hartley, M.D. Berardini, T.C. Jenkins, S. Neidle, L.H. Hurley, D.E. Thurston, Rational Design of a Highly Efficient Irreversible DNA Interstrand Cross-Linking Agent Based on the Pyrrolobenzodiazepine Ring-System, J. Am. Chem. Soc., 114 (1992) 4939-4941.
[21] G.P. Wilkinson, P.M. Loadman, J.P. Taylor, T.C. Jenkins, J.A. Double, S.J. Gregson, P.W. Howard, D.E. Thurston, Intracellular and in vivo distribution of the pyrrolobenzodiazepine dimer SJG-136, a novel sequence-selective DNA minor groove cross-linking agent, European Journal of Cancer, 38 (2002) S28-S29.
[22] C. Martin, T. Ellis, C.J. McGurk, T.C. Jenkins, J.A. Hartley, M.J. Waring, D.E. Thurston, Sequence-selective interaction of the minor-groove interstrand cross-linking agent SJG-136 with naked and cellular DNA: footprinting and enzyme inhibition studies, Biochemistry, 44 (2005) 4135-4147.
[23] S.R. Hopton, A.S. Thompson, Nuclear magnetic resonance solution structures of inter- and intrastrand adducts of DNA cross-linker SJG-136, Biochemistry, 50 (2011) 4720-4732.
[24] M.S. Kung Sutherland, R.B. Walter, S.C. Jeffrey, P.J. Burke, C. Yu, H. Kostner, I. Stone, M.C. Ryan, D. Sussman, R.P. Lyon, W. Zeng, K.H. Harrington, K. Klussman, L. Westendorf, D. Meyer, I.D. Bernstein, P.D. Senter, D.R. Benjamin, J.G. Drachman, J.A. McEarchern, SGN-CD33A: a novel CD33-targeting antibody-drug conjugate using a pyrrolobenzodiazepine dimer is active in models of drug-resistant AML, Blood, 122 (2013) 1455-1463.
[25] L.R. Saunders, A.J. Bankovich, W.C. Anderson, M.A. Aujay, S. Bheddah, K. Black, R. Desai, P.A. Escarpe, J. Hampl, A. Laysang, D. Liu, J. Lopez-Molina, M. Milton, A. Park, M.A. Pysz, H. Shao, B. Slingerland, M. Torgov, S.A. Williams, O. Foord, P. Howard, J. Jassem, A. Badzio, P. Czapiewski, D.H. Harpole, A. Dowlati, P.P. Massion, W.D. Travis, M.C. Pietanza, J.T. Poirier, C.M. Rudin, R.A. Stull, S.J. Dylla, A DLL3-targeted antibody-drug conjugate eradicates high-grade pulmonary neuroendocrine tumor-initiating cells in vivo, Science Translational Medicine, 7 (2015) 302ra136-302ra136.
[26] K.M. Rahman, H. Vassoler, C.H. James, D.E. Thurston, DNA Sequence Preference and Adduct Orientation of Pyrrolo[2,1-c][1,4]benzodiazepine Antitumor Agents, Acs Medicinal Chemistry Letters, 1 (2010) 427-432.
[27] J. Mantaj, Jackson, P. J. M., Rahman, K. M., Thurston, D. E., Interaction of SJG-136 with cognate sequences of oncogenic transcription factors, American Association of Cancer Research Annual MeetingWashington, 2013.
[28] P.J. Jackson, C.H. James, T.C. Jenkins, K.M. Rahman, D.E. Thurston, Computational Studies Support the Role of the C7-Sibirosamine Sugar of the Pyrrolobenzodiazepine (PBD) Sibiromycin in Transcription Factor Inhibition, ACS Chem Biol, 9 (2014) 2432-2440.
[29] M.S. Puvvada, J.A. Hartley, T.C. Jenkins, D.E. Thurston, A quantitative assay to measure the relative DNA-binding affinity of pyrrolo[2,1-c] [1,4]benzodiazepine (PBD) antitumour antibiotics based on the inhibition of restriction endonuclease BamHI, Nucleic Acids Res, 21 (1993) 3671-3675.
[30] P.H. Clingen, I.U. De Silva, P.J. McHugh, F.J. Ghadessy, M.J. Tilby, D.E. Thurston, J.A. Hartley, The XPF-ERCC1 endonuclease and homologous recombination contribute to the repair of minor groove DNA interstrand crosslinks in mammalian cells produced by the pyrrolo[2,1-c][1,4]benzodiazepine dimer SJG-136, Nucleic Acids Res, 33 (2005) 3283-3291.
[31] M.S. Puvvada, S.A. Forrow, J.A. Hartley, P. Stephenson, I. Gibson, T.C. Jenkins, D.E. Thurston, Inhibition of Bacteriophage T7 RNA Polymerase *In Vitro* Transcription by DNA-Binding Pyrrolo[2,1-*c*][1,4]benzodiazepines, Biochemistry, 36 (1997) 2478-2484.
[32] M. Ichimura, T. Ogawa, S. Katsumata, K. Takahashi, I. Takahashi, H. Nakano, Duocarmycins, new antitumor antibiotics produced by Streptomyces; producing organisms and improved production, The Journal of antibiotics, 44 (1991) 1045-1053.
[33] M. Ichimura, T. Ogawa, K. Takahashi, E. Kobayashi, I. Kawamoto, T. Yasuzawa, I. Takahashi, H. Nakano, Duocarmycin SA, a new antitumor antibiotic from Streptomyces sp, The Journal of antibiotics, 43 (1990) 1037-1038.
[34] T. Yasuzawa, T. Iida, K. Muroi, M. Ichimura, K. Takahashi, H. Sano, Structures of duocarmycins, novel antitumor antibiotics produced by Streptomyces sp, Chemical & pharmaceutical bulletin, 36 (1988) 3728-3731.
[35] W.H. Dokter, M.M. van der Lee, P.G. Groothuis, R. Ubink, M.A. van der Vleuten, T.A. van Achterberg, E.M. Loosveld, D. Damming, D.C. Jacobs, M. Rouwette, D.F. Egging, D. van den Dobbelsteen, P.H. Beusker, P. Goedings, G.F. Verheijden, J.M. Lemmens, M. Timmers, The preclinical profile of the duocarmycin-based HER2-targeting ADC SYD985 predicts for clinical benefit in low HER2-expressing breast cancers, Molecular cancer therapeutics, (2015).
[36] D.L. Boger, D.S. Johnson, W. Yun, C.M. Tarby, Molecular basis for sequence selective DNA alkylation by (+)- and ent-(-)-CC-1065 and related agents: alkylation site models that accommodate the offset AT-rich adenine N3 alkylation selectivity, Bioorganic & medicinal chemistry, 2 (1994) 115-135.
[37] A.R. Urbach, P.B. Dervan, Toward rules for 1:1 polyamide:DNA recognition, Proc Natl Acad Sci U S A, 98 (2001) 4343-4348.
[38] D.L. Boger, The Duocarmycins: Synthetic and Mechanistic Studies, Accounts of Chemical Research, 28 (1995) 20-29.
[39] D.L. Boger, D.S. Johnson, CC-1065 and the Duocarmycins: Understanding their Biological Function through Mechanistic Studies, Angewandte Chemie International Edition in English, 35 (1996) 1438-1474.
[40] D.L. Boger, R.M. Garbaccio, Catalysis of the CC-1065 and duocarmycin DNA alkylation reaction: DNA binding induced conformational change in the agent results in activation, Bioorganic & medicinal chemistry, 5 (1997) 263-276.
[41] D.L. Boger, R.M. Garbaccio, Are the Duocarmycin and CC-1065 DNA Alkylation Reactions Acid-Catalyzed? Solvolysis pH-Rate Profiles Suggest They Are Not, The Journal of organic chemistry, 64 (1999) 5666-5669.
[42] S.E. Wolkenberg, D.L. Boger, Mechanisms of in situ activation for DNA-targeting antitumor agents, Chem Rev, 102 (2002) 2477-2495.
[43] M.A. Mitchell, R.C. Kelly, N.A. Wicnienski, N.T. Hatzenbuhler, M.G. Williams, G.L. Petzold, J.L. Slightom, D.R.

Siemieniak, Synthesis and DNA crosslinking by a rigid CPI dimer, J. Am. Chem. Soc., 113 (1991) 8994-8995.
[44] G. Jia, J.W. Lown, Design, synthesis and cytotoxicity evaluation of 1-chloromethyl-5-hydroxy-1,2-dihydro-3H-benz[e]indole (seco-CBI) dimers, Bioorganic & medicinal chemistry, 8 (2000) 1607-1617.
[45] Q. Zhou, W.H. Duan, D. Simmons, Y. Shayo, M.A. Raymond, R.T. Dorr, L.H. Hurley, Design and Synthesis of a Novel DNA-DNA Interstrand Adenine-Guanine Cross-Linking Agent, J. Am. Chem. Soc., 123 (2001) 4865-4866.
[46] M. Tercel, S.M. Stribbling, H. Sheppard, B.G. Siim, K. Wu, S.M. Pullen, K.J. Botting, W.R. Wilson, W.A. Denny, Unsymmetrical DNA Cross-Linking Agents: Combination of the CBI and PBD Pharmacophores, J. Med. Chem., 46 (2003) 2132-2151.
[47] B. Purnell, A. Sato, A. O'Kelley, C. Price, K. Summerville, S. Hudson, C. O'Hare, K. Kiakos, T. Asao, M. Lee, J.A. Hartley, DNA Interstrand Crosslinking Agents: Synthesis, DNA interactions, and Cytotoxicity of Dimeric Achiral Seco-Amino-CBI and Conjugates of Achiral Seco-Amino-CBI with Pyrrolobenzodiazepine (PBD), Bioorganic & Medicinal Chemistry Letters, 16 (2006) 5677-5681.
[48] P.J.M. Jackson, Development of In Silico Techniques for the Rational Design of Novel Pyrrolobenzodiazepine-Based Transcription Factor Inhibitors, Institute of Pharmaceutical Science, King's College London, London, 2014.
[49] B. Purnell, A. Sato, A. O'Kelley, C. Price, K. Summerville, S. Hudson, C. O'Hare, K. Kiakos, T. Asao, M. Lee, J.A. Hartley, DNA interstrand crosslinking agents: Synthesis, DNA interactions, and cytotoxicity of dimeric achiral seco-amino-CBI and conjugates of achiral seco-amino-CBI with pyrrolobenzodiazepine (PBD), Bioorganic & Medicinal Chemistry Letters, 16 (2006) 5677-5681.
[50] J.A. Hartley, V.J. Spanswick, N. Brooks, P.H. Clingen, P.J. McHugh, D. Hochhauser, R.B. Pedley, L.R. Kelland, M.C. Alley, R. Schultz, M.G. Hollingshead, K.M. Schweikart, J.E. Tomaszewski, E.A. Sausville, S.J. Gregson, P.W. Howard, D.E. Thurston, SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity: part 1: cellular pharmacology, in vitro and initial in vivo antitumor activity, Cancer Res, 64 (2004) 6693-6699.
[51] S.J. Gregson, P.W. Howard, D.R. Gullick, A. Hamaguchi, K.E. Corcoran, N.A. Brooks, J.A. Hartley, T.C. Jenkins, S. Patel, M.J. Guille, D.E. Thurston, Linker length modulates DNA cross-linking reactivity and cytotoxic potency of C8/C8' ether-linked C2-exo-unsaturated pyrrolo[2,1-c][1,4]benzodiazepine (PBD) dimers, J Med Chem, 47 (2004) 1161-1174.


Compound 11
(blue)
A3
(yellow)
G12
(magenta)


5'-A3-T4-T5-T6-T7-C8-C9 -3'
3'-T18-A17-A16-A15-A14-G13-G12-5'