MedChemComm

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/medchemcomm

MedChemComm

COMMUNICATION

Improved Antiviral Activity of a Polyamide Against High-Risk Human Papillomavirus *Via* N-Terminal Guanidinium Substitution§

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/medchemcomm

C. H. Castaneda,[†a] M. J. Scuderi,[†a] T. G. Edwards,[‡a] G. D. Harris Jr.,[‡a] C. M. Dupureur,[*a] K. J. Koeller,[*a] C. Fisher,[*b] J. K. Bashkin[*a,b]

[a.] *Department of Chemistry & Biochemistry, University of Missouri–St. Louis, St. Louis, MO 63121, USA.*
[b.] *NanoVir, LLC, Kalamazoo, MI 49008 (USA).*
† These authors contributed equally to this work.
‡ These authors contributed equally to this work.
§ The authors declare no competing interests.
Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

MedChemComm

COMMUNICATION

We report the synthesis of two novel pyrrole-imidazole polyamides with N-terminal guanidinium or tetramethylguanidinium groups and evaluate their antiviral activity against three cancer-causing human papillomavirus strains. Introduction of guanidinium improves antiviral activity when compared to an unsubstituted analog, especially in IC_{90} values. These substitutions change DNA-binding preferences, while binding affinity remains unchanged.

N-methylpyrrole (Py) and N-methylimidazole (Im) hairpin polyamides (PAs) are synthetic homologs of the natural antibiotics distamycin A and netropsin which bind AT-rich regions of DNA in the minor groove.[1] The antiparallel nature of hairpin PAs allows for the side-by-side pairing of building blocks and recognition of DNA base pairs (bp) within a double strand (dsDNA). It has been extensively reported that Py/Py and Py/β-alanine (β) pairs recognize A/T and T/A base pairs, while an Im/Py pair distinguishes G/C from C/G.[2] Furthermore, introduction of the flexible β linker in place of heterocycles resets the heterocycle-base pair register and allows targeting of longer DNA sequences.[3]

In numerous reports hairpin PAs distinguish DNA sequences by binding to patterns of hydrogen-bond donors and acceptors in the minor groove of DNA,[4] so synthetic PAs are attractive for many biomedical applications. Thus, for polyamides themselves and related minor groove binders, reports include the regulation of gene expression,[5] biological imaging,[6] and use as antimicrobial/antiviral agents.[7] Antiviral polyamides include the long polyamides reported here which, as a class, access a new biological mechanism: they lead to the degradation of the genome of dsDNA tumour viruses in cells while avoiding acute cytotoxicity. The DNA degradation is accomplished *via* the DNA Damage Response (DDR), for which expression is altered significantly only in infected cells carrying small circular, dsDNA genomes of the virus, known as episomes.[8a,b] As discussed below, another characteristic of this new class of antiviral polyamides is their failure to follow the reported DNA recognition rules described above.

The process by which the DDR works in concert with the polyamides is not completely established at the molecular level, but many details have emerged as a result of a variety of experiments, a portion of which are summarized here. For example, 224 DDR genes were knocked down individually using siRNA, and we checked to see if the knockdowns affected polyamide activity in any way; 21 genes were found to affect PA activity, and the results were validated by additional experiments using QPCR (rather than relying on the PCR array data alone).[8a,b] Some gene knockdowns enhanced PA activity and others opposed PA activity. Additional experiments used Mirin, a small molecule inhibitor of DDR enzyme Mre11. Mre11 has many activities alone and in protein complexes. For example, Mre11 is a member of the MRN complex, which participates in double-strand DNA break repair and telomere maintenance as a member of the ATM branch of the DDR.[8c] As a standalone enzyme, one activity associated with MRE11 is that of an endonuclease. MRE11 inhibition experiments were carried out in the presence and absence of polyamides.[8a] 100 μM Mirin had no effect on its own, but it sensitized cells to PA25, making the compound 4x more active in terms of lower IC_{50} and IC_{90} values. Part of the activation of the DDR was also followed by phosphorylation of Rad9, a protein that is phosphorylated upon DNA damage as part of the ATR pathway of the DDR, one of the two major DDR branches (the other being ATM).[8a,b,d] Treatment of HPV16 episome bearing W12E cells with active polyamide causes Rad9 phosphorylation within 2-4 hours, but HPV-free cervical cancer cell line C33A shows no phosphorylation of Rad9 upon treatment with polyamide.[8a,b]

We previously reported a library of long hairpin PAs, including **PA1** and **PA25**, which exhibits antiviral activity against three high-risk (oncogenic) HPVs.[9] HPV is an AT-rich dsDNA virus[10] that causes cervical cancer, among other diseases, and subtypes HPV16 and 18 are responsible for 70% of cervical cancer in most of the world.[11,12] Although specific HPV vaccines are available,[13] they are effective prophylactics, not therapies. **PA1**, our first preclinical lead, exhibits anti-HPV activity with no cytotoxicity up to its solubility limit in cell culture media of 200 μM.[9] A characteristic of these active anti-HPV polyamides is that they bind at least ten bp or one full turn of DNA.

Here we report the antiviral activities and DNA-binding properties of two PAs of a new structural class, **PA30**, and **PA31**, where the des-amino N-terminus of the parent compound **PA1** has been substituted with either tetramethylguanidinium (TMG) or guanidinium (Guan), respectively. **PA30** and **PA31** are biomimetic in that netropsin has an N-terminal guanidinium, though **PA30**, **31** are not exact structural analogues of netropsin. While others have explored guanidinylated DNA-binding ligands, no biophysical studies on guanidinylated hairpin PAs have been undertaken, to our knowledge.[14] We found that these substitutions at the N-terminus, which are small relative to the overall molecular size, can significantly improve anti-HPV activity. However, improvement in antiviral efficacy cannot be attributed to altered DNA binding strength, as will be shown. Nevertheless, **PA30** and **PA31** display noticeable differences in binding distribution on a fragment of native HPV18 DNA in comparison with parent **PA1**.

Table 1 IC_{50} and IC_{90} values of **PA1**,[a] **PA30,** and **PA31** against high-risk HPV16, -18 and -31 types.

	HPV16			HPV18			HPV31		
	IC_{50} (μM)	IC_{90} (μM)	n	IC_{50} (μM)	IC_{90} (μM)	n	IC_{50} (μM)	IC_{90} (μM)	n
PA1	0.1(1)	1.1(8)	4	0.7(4)	>10	3	0.1(1)	1.0(5)	4
PA30	0.3(1)	>10	3	0.2(2)	>10	3	0.13(4)	0.9(7)	3
PA31	0.10(1)	0.38(8)	3	0.2(2)	0.9(6)	3	0.14(4)	0.8(2)	3

The numbers in parentheses are sample standard deviations; n is the number of independent measurements. [a] IC_{50} values for **PA1** were first reported in reference 9; we re-measured them and found them unchanged.

Polyamides were synthesized by Boc solid-phase methods[15] as reported[9] (**PA1**) or followed by TMG (**PA30**) or Guan (**PA31**) conjugation. After cleavage from the resin, polyamides were purified by reverse-phase HPLC and characterized with ^{1}H and ^{13}C NMR and HR mass spectrometry (SI). Structures of **PA1**, **PA30** and **PA31** and their EDTA-conjugates used in affinity cleavage maps are shown in Figure 1. Anti-HPV activities and MTT cell toxicity assays (shown for **PA1**, **Fig. SI3A**) were determined as previously described.[9a] New side-by-side, concentration-dependent assays were performed for previously reported **PA1**.[9a] Table 1 provides the resulting pseudo IC_{50} and IC_{90} values for **PA1**, **PA30**, and **PA31** corresponding to the PA concentration that causes a 50% or 90% decrease in viral DNA concentration. **Figure SI2** gives dose-response curves for **PA30** and **31** while data are given for **PA1** and HPV16, 31 in **Figure SI3**; data for **PA1** were the same as previously reported.[9a,26] Although IC_{50} values for the compounds are comparable against HPV16 and 31, **PA30** and **PA31** afforded somewhat improved IC_{50} values against HPV18. In contrast, significant improvement in IC_{90} for HPV16 and -18 was observed with **PA31**. However, **PA30** activity plateaus just below a 90% decrease in viral DNA and, in two days, does not quite reach an IC_{90}, an important measure of antiviral activity. We will explore longer time intervals, but higher concentrations do not seem promising based on the dose-response curves, and would make **PA30** a poor competitor with the more-active **PA31**.

Because only **PA31** showed an improved IC_{90} against HPV18, we investigated the DNA-binding properties of these hairpin PAs on a natural HPV18 sequence and compared them to the parent compound, **PA1**. Based on well-established PA-DNA pairing rules,[4a] **PA1**, **PA30**, and **PA31** were expected to target the ten base-pair DNA sequence WWGWWWWWWW, where W=A or T (Figure 1 inset). A PA-DNA mismatch refers to any deviation from this recognition motif, i.e. where a PA building block pair binds to a noncanonical Watson-Crick base pair: one not predicted by the pairing rules.[1a] We used quantitative DNase I footprinting[16] to obtain dissociation constants (K_d) and affinity cleavage experiments to assign binding orientations of the compounds on a 305 bp DNA fragment corresponding to nucleotides (nt) 7479 to 7783 in the Long Control Region of HPV18, a viral genomic region containing regulatory elements for replication.[17] We determined K_d values and orientation for all PA binding events along the entire 305 bp DNA fragment, and representative dissociation constants at three binding sites are given in Table 2. Similar results were observed in the rest of the DNA fragment. K_d was determined by the Hill or Langmuir equation, based on the magnitude of the derived Hill coefficient and overall fit (see Table S 1 and 2).[18]

There is no significant difference in DNA-binding affinities of **PA1**, **PA30**, and **PA31**, indicating that the addition of a potential H-bond donor and a positive charge for **PA30** and **PA31** *vs.* **PA1** does not explain enhanced antiviral activity. It is noteworthy that since we are sampling an ensemble of different binding conformations in multiple linear, cell-free DNA molecules, binding affinities are the sum of these binding events. Furthermore the negative supercoiling of the viral genome, and bound viral and host proteins *in vivo,* probably affect binding (see Conclusions). We are currently examining these matters experimentally. All three PAs bind to Sites 1 and 2 (single-mismatch sites) with K_d values ranging from 0.9-1.2 nM, and it is noteworthy that they all tolerate a triple base-pair mismatch at Site 3 with K_d values ranging from 2.2-3.2 nM (Table 2). The high Hill coefficients at this site are explained by previously-reported polyamide binding cooperativity,[4d] in which the binding of one polyamide can preorganize the minor groove for the binding of a second (and more) polyamides, for example by widening the minor groove. This cooperativity would be most expected in areas of high polyamide density on the DNA, which is what we found in the long control region (LCR). We need to point out that we used a single-site binding model to analyze binding thermodynamics. If we had parameterized binding to allow multiple Hill coefficients to float simultaneously, we would more accurately have accounted for the statistics of multi-site binding. The results would likely have been that some Hill Kd values and coefficients would have decreased, and others would have increased. However, we were not convinced that our data would hold up to the introduction of so many additional parameters (a Hill coefficient per binding site and orientation), and did not believe that the outcome would change significantly in a scientific sense with a more precise handling of the statistical question. We will revisit this with more complete data sets for multiple compounds to test whether our approximation, which is normal in polyamide literature,[1-4] is suitable or not.

Thus, large antiviral hairpin PAs tolerate reportedly unfavourable DNA interactions without a significant decrease in binding affinity, which agrees with our prior work on HPV16.[18] In fact, in that prior work we found for 14-ring **PA1** that numerous "mismatch" sites were bound with higher affinity than so-called canonical sites. Also, up to four mismatches were tolerated with 2 nM dissociation constants identical to numerous "perfect-match" sites. For a larger, 20-ring antiviral polyamide (**PA25**), four mismatches were even more common than with **PA1**, leading to a complete blanketing of the HPV16 long control region when only sparse binding, at best, was predicted.[18b]

Table 2 Representative PA binding sites on nt 7479-7783 of the HPV18 LCR as determined by quantitative DNase I footprinting and affinity cleavage.

				K_d (nM)		
Site	**Sequence**	**Position**	**Site Type**	**PA1**	**PA30**	**PA31**
1	CCTGG TATTA**G**T<u>C</u>AT TTTCC	7606-7615	Single-mismatch	0.9 ± 0.1	1.2 ± 0.2	1.0 ± 0.1
2	ACATA TTTT**G**AA<u>C</u>AA TTGGC	7559-7568	Single-mismatch	0.9 ± 0.2	1.1 ± 0.2	1.0 ± 0.1
3	CTTTG **GC**<u>G</u>**C**ATATAA GGCGC	7585-7594	Triple-mismatch	2.7 ± 0.5	3.2 ± 0.3	2.2 ± 0.2

DNA sequences are illustrated so that the middle block represents the PA binding site along with 5 flanking nt. In each sequence, the underlined base is bound by imidazole/pyrrole (or pyrrole/imidazole) pair of the polyamide. The bold bases in column 2 correspond to mismatches according to recognition "rules." K_d values were calculated from DNase I experiments performed at least in triplicate (Table S 1 and 2).

Data for HPV18-polyamide binding to a largely complete LCR is reported here and allows us to expand our understanding (polyamide-HPV18 binding was previously reported for the E2-protein binding sites).[7a] We hypothesize that a deviation from established binding rules occurs for our unusually large hairpin PAs because they form many H-bonds with DNA and have significantly more internal π-stacking than the smaller PAs typically found in the literature.[19] Thus, a few unfavourable enthalpic interactions are likely insignificant. The entropy of binding is likely very favourable for long PAs from the loss of many water molecules previously resident in the minor groove and from loss of exposed nonpolar surface area upon π-stacking. Entropy and enthalpy are being determined by the CM Dupureur group. Since we measured polyamide-DNA binding properties in the context of a large, natural DNA sequences instead of small artificial, DNA sequences designed to examine sequence preferences at a single location, our results are much more complex than those found in the literature. Published binding rules are not obeyed in our system with significant fidelity, and "mismatches" are very well tolerated (see also He et al.[18] and Qiao et al.[20]).

Binding sites for **PA1**, **PA30**, and **PA31** in a region of HPV18DNA spanning nt 7601-7620 are shown in Figure 2A. In this part of the viral genome, there are two predicted single-mismatch binding sites, sequence **(i)** 5'-TATTAGTCAT-3' (7606-7615) and sequence **(ii)** 5'-TAGTCATTTT-3' (7609-7618). Our data indicate that **PA1** and **PA31** bind both of these sites, while **PA30** only binds to **(i)**.

Raw DNase I footprinting data of dsDNA (top strand) is shown in Figures 2B (control) and 2C (5 nM PA). Inspection of the footprints generated by the three PAs reveals subtle differences in the protected regions. For instance, the peaks at the left-hand edge of the protected region corresponding to nt 7606-09 (Figure 2* shows 7606) display increased nuclease accessibility in the case of **PA1** vs. **PA30** and **PA31**. At the right-hand edge of the protected region, the peak corresponding to nt 7619 (Figure 2C[†]) is less protected from DNase I by **PA30** (Figure 2C middle panel) than by **PA1** (Figure 2C left panel) at equivalent PA concentrations. In contrast, **PA31** affords more-extended nuclease protection than either **PA1** or **PA30**, with a footprint including extra nt 7619-7626; this extra nuclease protection corresponds to the double-mismatch site **(iii)** at 7612-7621. For steric reasons, the TMG group of **PA30** may not bind DNA, decreasing productive enthalpic interactions compared to **PA1** and **PA31**.

Footprinting differences are supported by affinity cleavage (AC) results.[21] Patterns generated by **PA1-**, **PA30-** and **PA31-EDTA** conjugates are given in Figure 2D. For **PA1** and **PA31** there are two binding events, each corresponding to a distinct AC pattern, agreeing with DNase I data that both **PAs** bind at both available single-mismatch sites, **(i)** and **(ii)**. However, an extra AC pattern was observed for PA31 corresponding to the double mismatch site **(iii)**. Conversely, only one of the predicted single-mismatch sites is occupied by **PA30**: a single AC pattern corresponding to site **(i)** was observed. The AC pattern for site **(ii)** did not appear even upon increasing the concentration of **PA30** from 50 nM to 200 nM (data not shown). Our results with these 14-ring PAs agree with previous 8-ring PAs studies reporting that substitution at the N-terminal imidazole can modulate orientation and DNA-binding preferences of hairpin polyamides.[22]

Conclusions

In summary, we show that N-terminal substitution of anti-HPV polyamide **PA1** improves antiviral activity against the two most high-risk HPV types that are most prevalent (HPV16 and -18). N-terminal substitution with TMG gives **PA30** and leads to improvement in anti-HPV activity (IC_{50}) against HPV18. N-terminal substitution with guanidinium itself gives **PA31** with enhanced anti-HPV activity even further relative to **PA1** by decreasing IC_{50} against HPV18 and IC_{90} for both HPV16 and 18. For antiviral agents in general, a dosage above IC_{90} decreases viral load, magnitude and frequency of viral rebound,[23] and the probability of viral mutation, so improvements in IC_{90} are significant for therapy, as they are for EC_{90}.[24]

Although the dissociation constants for the reported PAs are similar and cannot account for the differences in antiviral activity, differences in PA-DNA binding distribution were observed: **PA31** binds more sites on the LCR of HPV18 than **PA1** or **PA30**. Binding kinetics may be important for activity[25] and remain under investigation. Polyamide uptake is also under investigation. We note that **PA1** and its analogs **PA30** and **PA31** are among the smallest molecules in a library of 75+ active compounds that reach nearly twice the molecular weight of **PA1**. We have also prepared a series of compounds based on the highly-active **PA25** and other active series, and they will be described subsequently. **PA25** was our second preclinical lead because of its IC_{50}/IC_{90} data (μM) as follows: HPV16, 0.036/0.35; HPV18, 0.056/1.5; HPV31, 0.030/0.51.[9,26] The first biophysical studies of **PA25** were reported recently.[18b]

However, **PA25** is a more difficult synthetic and purification challenge than **PA1** and its derivatives, though we have scaled up both **PA1** and **PA25** to 10 g, and the cost-of-goods is much higher for the 20-ring **PA25** than the 14-ring **PA1**, so we were interested to see if modifications to **PA1** could increase its activity significantly. That goal was successful with **PA31**.

Long hairpin polyamides of this general class function by a novel polyamide mechanism that invokes the host DNA Damage Response to destroy the viral genome.[8,10] Their activity is not limited to HPV but is also found for other small, double-stranded DNA tumour viruses, including polyomaviruses.[26] This mechanism is unrelated to any prior polyamide literature that we can determine, even though our work was initially inspired by early PA research. **PA30/31** represent an improved class of long polyamide that invokes this new mechanism to combat cancer-causing HPV. More recently, Dervan has reported a similar, non-sequence-specific, polyamide-induced DNA Damage Response.[27]

Why might this nonspecific binding, or poorly-specific binding, be so relevant? We have come to hypothesize and are performing experiments to test the idea, that the interaction of large polyamides with small, circular, double-stranded DNA genomes is biologically important more from physical than chemical reasons. Hydrogen bonding does go on in the minor groove, albeit at moderate selectivity, but of great importance is the binding of many hundreds or more of rigid, crescent-shaped molecules to supercoiled DNA targets. This must result in considerable strain on the supercoiled viral genomes, and could in the most extreme case result in single-stranded DNA regions appearing, which would be rapidly noticed and acted upon by the DNA damage response machinery. However, the process may be simpler: the dsDNA viruses evolved to use parts of the DDR for their life cycles but otherwise to hide, to avoid having their genomes recognized as foreign or damaged in any way. When the supercoiled genomes are distorted by the effect of many bound polyamides, and 50% of HPV DNA is bound by active polyamides, at least for the linearized pieces we have reported on to date,[9c,18] the viral genome may no longer remain hidden. The strain imposed by all of the bound polyamides may change the shape enough so that the DDR naturally notices unwanted additions to the chromatin and is automatically activated. Support for this hypothesis is found in our recently published patent application, where, as mentioned above, we showed that the same polyamides active against HPV are also active against three polyomaviruses having little in common with HPV except a small, dsDNA genome that is supercoiled.[26] The DNA damage response is also activated for these disparate viruses.

Notes and references

Declaration: JKB and CF hold significant equity positions in NanoVir.

We thank Jacqui Niederschulte for extinction coefficient determination of **PA1**, the DNA Core at the University of Missouri for DNA CE fragment analysis, the Danforth Plant Science Center (NSF DBI 0922879], Prof. B. Bythell for HRMS, the referees for helpful comments, and Prof. P. Lambert for providing W12E cells. The Agilent 600 MHz NMR spectrometer was obtained using funds from NSF (MRF #0959360). This work was supported by NIH-NIAID: AI083803, AI062182, and AI068159.

1 (*a*) P. B. Dervan, B. S. Edelson, *Curr. Opin. Struct. Biol.*, 2003, **13**, 284; (*b*) X. Chen, B. Ramakrishnan, S.T. Rao, M. Sundaralingam, *Nat. Struct. Mol. Biol.*, 1994, **1**, 169; (*c*) L. A. Marky, K. J. Breslauer, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 4359; (*d*) A. C. Finlay, F. A. Hochstein, B. A. Sobin, F. X. Murphy, *J. Am. Chem. Soc.*, 1951, **73**, 341.

2 *(a)* S. White, E. E. Baird, P. B. Dervan, *Biochemistry*, 1996, **35**, 12532. *(b)* S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature*, 1998, **391**, 468-471.

3 (*a*) J. W. Trauger, E. E. Baird, M. Mrksich, P. B. Dervan, *JACS* 1996, **118**, 6160; (*b*) C.C.C. Wang, U. Ellervik; P. B. Dervan, *Bioorg. Med. Chem.*, 2001, **9**, 653.

4 (*a*) S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature*, 1998, **391**, 468; (*b*) M. Mrksich, W. S. Wade, T. J. Dwyer, B. H. Geierstanger, D. E. Wemmer, P. B. Dervan, *Proc. Natl. Acad. Sci., USA* 1992, **89**, 7586; (*c*) M. Mrksich, M. E. Parks, P. B. Dervan, *J. Am. Chem. Soc.*, 1994, **116**, 7983; *(d)* C. A. Hawkins, E. E. Baird, P. B. Dervan and D. E. Wemmer, *J. Am. Chem. Soc.*, 2002, **124**, 12689.

5 (*a*) A. K. Mapp, A. Z. Ansari, M. Ptashne, P. B. Dervan, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 3930; (*b*) X. Wang, H. Nagase, T. Watanabe, H. Nobusue, T. Suzuki, Y. Asami, Y. Shinojima, H. Kawashima, K. Takagi, R. Mishra, J. Igarashi, M. Kimura, T. Takayama, N. Fukuda, H. Sugiyama, *Cancer. Sci.*, 2010, **101**, 759; (*c*) J. A. Raskatov, J. L. Meier, J. W. Puckett, F. Yang, P. Ramakrishnan, P. B. Dervan, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 1023.

6 K. Maeshima, S. Janssen, U. K. Laemmli, *EMBO J.*, 2001, **20**, 3218.

7 (*a*) T. D. Schaal, W. G. Mallet, D. L. McMinn, N. V. Nguyen, M. M. Sopko, S. John, B. S. Parekh, *Nucleic Acids Res.*, 2003, **31**, 1282; (*b*) A. Yasuda, K. Noguchi, M. Minoshima, G. Kashiwazaki, T. Kanda, K. Katayama, J. Mitsuhashi, T. Bando, H. Sugiyama, Y. Sugimoto, *Cancer Sci.*, 2011, **102**, 2221; (*c*) L. A. Dickinson, R. J. Gulizia, J. W. Trauger, E. E. Baird, D. E. Mosier, J. M. Gottesfeld, P. B. Dervan, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 12890; *(d)* R. W. Burli, D. McMinn, J. A. Kaizerman, W. Hu, Y. Ge, Q. Pack, V. Jiang, M. Gross, M. Garcia, R. Tanaka, H. E. Moser, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 1253-1257; *(e)* R. W. Burli, P. Jones, D. McMinn, Q. Le, J.-X. Duan, J. A. Kaizerman, S. Difuntorum, H. E. Moser, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 1259-1263; *(f)* M. P. Barrett, C. G. Gemmell, C. J. Suckling, *Pharmacol. Ther.*, 2013, **139**, 12-23; *(g)* M. N. C. Soeiro, K. Werbovetz, D. W. Boykin, W. D. Wilson, M. Z. Wang, A. Hemphill, *Parasitology*, 2013, **140**, 929-951.

8 (*a*) T. G. Edwards, T. J. Vidmar, K. Koeller, J. K. Bashkin, C. Fisher, *PLoS One*, 2013, **8**, e75406. (*b*) C. Fisher, *J. Clin. Med.*, 2015, **4**, 204. (*c*) B. J. Lamarche, N. I. Orazio, M. D. Weitzman, *FEBS Lett.,* 2010, **584**, 3682-3695; (*d*) J. J. Lin, A. Dutta, *J. Biol. Chem.,* 2007, **282**, 30357-30362.

9 *(a)* T. G. Edwards, K. J. Koeller, U. Slomczynska, K. Fok, M. Helmus, J. K. Bashkin, C. Fisher, *Antiviral Res.*, 2011, **91**, 177; *(b)* G. He, J. K. Bashkin, *Future Med. Chem.*, 2015, **7**, 1953-1955; *(c)* K. J. Koeller, G. D. Harris, K. Aston, G. He, C. H. Castaneda, M. A. Thornton, T. G. Edwards, S. Wang, R. Nanjunda, W. D. Wilson, C. Fisher, J. K. Bashkin, *Med Chem (Los Angeles)*, 2014, **4**, 338-344.

10 T. G. Edwards, M. J. Helmus, K. Koeller, J. K. Bashkin, C. Fisher, *J. Virol.*, 2013, **87**, 3979.

11 J. M. M. Walboomers, M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. F. Snijders, J. Peto, C. J. L. M. Meijer, N. Muñoz, *J. Pathol.*, 1999, **189**, 12.

12 G. Clifford, S. Franceschi, M. Diaz, N. Munoz, L. L. Villa, *Vaccine*, 2006, **24** *(Suppl 3)*, S26.

13 (*a*) C. Dochez, J. J. Bogers, R. Verhelst, H. Rees, *Vaccine*, 2014, **32**, 1595; (*b*) E. A. Joura, A. R. Giuliano, O.E. Iversen, C. Bouchard, C. Mao, J. Mehlsen, E. D. Moreira, Y. Ngan, L. K. Petersen, E. Lazcano-Ponce, P. Pitisuttithum, J. A. Restrepo, G. Stuart, L. Woelber, Y. C. Yang, J. Cuzick, S. M. Garland, W. Huh, S. K. Kjaer, O. M. Bautista, I. S. F. Chan, J. Chen, R. Gesser, E. Moeller, M. Ritter, S. Vuocolo, A. Luxembourg, *N. Engl. J. Med.*, 2015, **372**, 711.

14 (*a*) P. S. Nagle, C. McKeever, F. Rodriguez, B. Nguyen, W. D. Wilson, I. Rozas, *J. Med. Chem.*, 2014, **57**, 7663; (*b*) R. K. Arafa, M. A. Ismail, M. Munde, W. D. Wilson, T. Wenzler, R. Brun, D. W. Boykin, *Eur. J. Med. Chem.*, 2008, **43**, 2901; (*c*) F. Rodríguez, I. Rozas, M. Kaiser, R. Brun, B. Nguyen, W. D. Wilson, R. N. García, C. Dardonville, *J. Med. Chem.*, 2008, **51**, 909.

15 *(a)* E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.*, 1996, **118**, 6141. *(b)* J. W. Puckett, J. T. Green, P. B. Dervan, *Org. Lett.*, 2012, **14**, 2774-2777; *(c)* note: in the hands of a microwave peptide synthesis machine manufacturer over an extended period, microwave-assisted synthesis was not successful for **PA1**, failing after the γ turn.

16 (*a*) D. J. Galas, A. Schmitz, *Nucleic Acids Res.*, 1978, **5**, 3157; (*b*) W. Yindeeyoungyeon, M. A. Schell, *Biotechniques*, 2000, **29**, 1034.

17 W. Stünkel, H.-U. Bernard, *J. Virol.*, 1999, **73**, 1918.

18 *(a)* G. He, E. Vasilieva, G. D. Harris Jr, K. J. Koeller, J. K. Bashkin, C. M. Dupureur, *Biochimie*, 2014, **102**, 83; *(b)* E. Vasilieva, J. Niederschulte, Y. Song, J. George D Harris, K. J. Koeller, P. Liao, J. K. Bashkin, C. Dupureur, *Biochimie*, 2016, **127**, 103-114.

19 J. M. Turner, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.*, 1997, **119**, 7636.

20 H. Qiao, C. Ma, X. Zhang, X. Jing, C. Li, and Y. Zhao, *Bioconjugate Chem.*, 2015, **26,** 2054.

21 S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.*, 1996, **118**, 8198.

22 (*a*) S. White, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.*, 1997, **119**, 8756; (*b*) C. A. Hawkins, R. P. de Clairac, R. N. Dominey, E. E. Baird, S. White, P. B. Dervan, D. E. Wemmer, *J. Am. Chem. Soc.*, 2000, **122**, 5235.

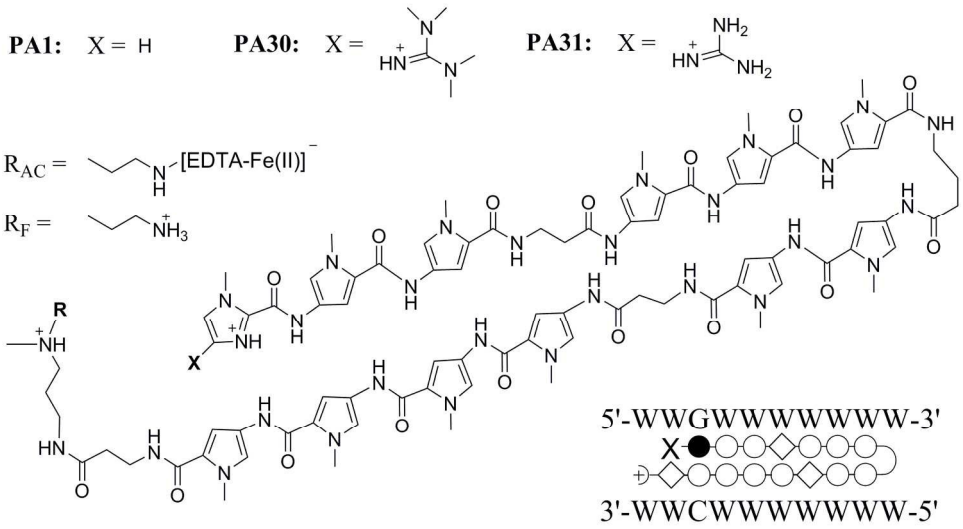
23 J. M. Pawlotsky, *Hepatology*, 2011, **53**, 1742.

24 S. Duffy, L. A. Shackelton, E. C. Holmes, *Nat. Rev. Genet.*, 2008, **9**, 267.

25 (*a*) M. Goyal, M. Rizzo, F. Schumacher, C. F. Wong, *J. Med. Chem.*, 2009, **52**, 5582; (*b*) R. A. Copeland, D. L. Pompliano, T. D. Meek, *Nat. Rev. Drug Discov.*, 2006, **5**, 730.

26 J. K. Bashkin, T. G. Edwards, C. Fisher, G. D. Harris, Jr., K. J. Koeller, US Patent Office, Patent number: 20150329596, Appln. Number: 14/818881, November 19, 2015.

27 T. F. Martinez, J. W. Phillips, K. K. Karanja, P. Polaczek, C. Wang, B. C. Li, J. L. Campbell, P. B. Dervan, *Nucleic Acids Res.*, 2014, **42**, 11546.



MTT cell toxicity assays (show
183x97mm (300 x 300 DPI)

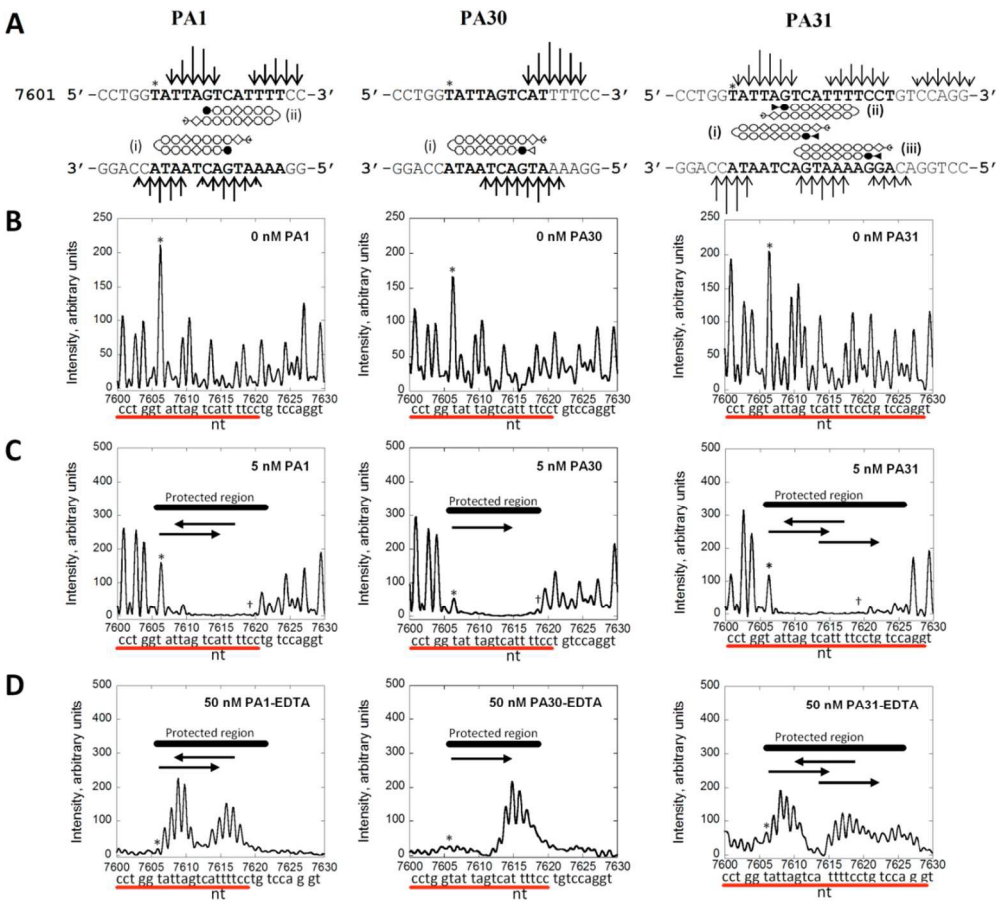
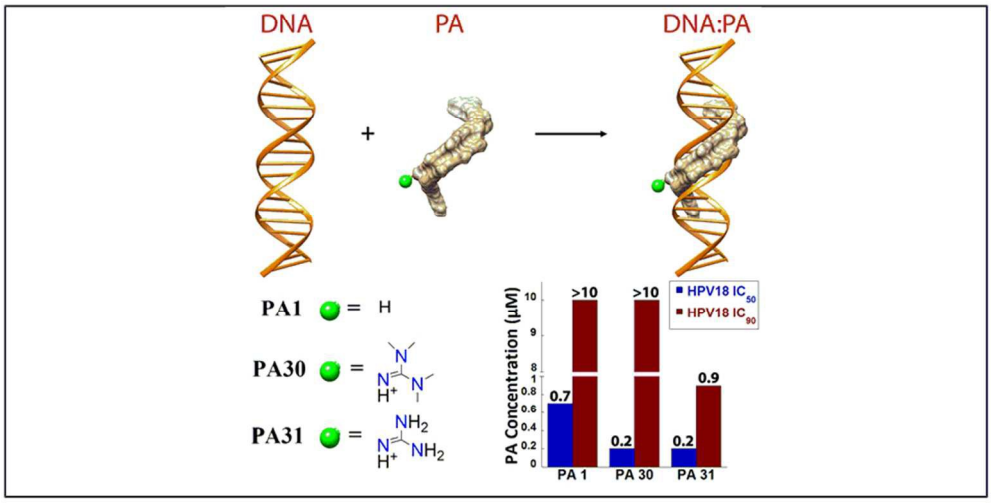


Fig. 2 (A) Binding sites for PA1, PA30, and PA31 from a section of HPV18 LCR corresponding to 7600-7620 as determined by capillary electrophoresis, DNase I footprinting and AC. The relative heights above the DNA sequence correspond to observed AC intensities. The remaining panels show the top strand of ten more base pairs than panel A, 7600–7630. Representative electropherograms showing the top strand for (B) DNase I footprinting controls (0 nM PA), (C) DNase I-protected regions upon the addition of 5 nM of PA and (D) AC patterns at 50 nM PA. The red line below each plot corresponds to the DNA sequence shown in panel (A). The * and † symbols mark nt 7606 and 7619. Arrows inside the plots show the PA binding direction, with the arrowhead in the position of the PA tail. PA building blocks as defined in Fig. 1; TMG and Guan groups are open and filled triangles.

Data for HPV18-polyamide bindi
147x132mm (220 x 220 DPI)

RSC Author Templates - ChemDraw (CDX) - Graphical Abstracts

All text and images must be placed within the frame.



81x71mm (300 x 300 DPI)