

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

C. H. Castaneda, M. J. Scuderi, T. G. Edwards, G. D. Harris Jr., C. M. Dupureur, K. J. Koeller, C. Fisher, J. K. Bashkin

Supporting Information

Contents

- I. Polyamide Synthesis and Characterization
- II. Compound Efficacy Testing and Results
- III. DNase I Footprinting
- IV. Affinity Cleavage

I. Polyamides Synthesis and Characterization

PA1 and its EDTA conjugate were synthesized and characterized as previously described.^[9, 18] **PA30** and **31** were synthesized by virtually identical solid-phase methods up to the Im group (4-amino-N-methyl imidazole-2-carboxamide) and its 4-N-attached guanidinyll substituent. For **PA30**, HATU^[S11] was utilized as the final building block to incorporate the TMG group onto the H₂N-Im moiety. For **PA31**, in parallel to the literature,^[S12] treatment with N,N'-Bis (tert-butoxycarbonyl)-1H-pyrazole-1-carboxamidine and diisopropylethylamine for 24 h, followed by TFA treatment to remove the BOC groups produced the desired, unsubstituted Guan group. The EDTA conjugate of **PA30** was prepared via the same method as previously described for **PA1**'s conjugate.^[18] Briefly, to a vigorously-stirred mixture of EDTA dianhydride (15.8 mg, 0.060 mmol, 22 eq) in N,N-diisopropylethylamine (DIEA) (0.5 mL), DMF (0.25 mL) and DMSO (0.25 mL) at 55 °C was added a mixture of **PA30** (5.9 mg, 0.0027 mmol, 1 eq) in DIEA (0.5 mL) and DMF (0.5 mL) dropwise over a period of 30 min. The reaction mixture was stirred at 55 °C for 30 min. Following addition of 0.1 N NaOH (1.03 mL), the mixture was stirred at 55 °C for an additional 20 min. The bottom layer of the biphasic mixture was removed from the reaction vessel and neutralized with formic acid (0.013 mL). The mixture was diluted with 0.5 mL DMSO, filtered through a 20 µm polyethylene filter, and purified by reversed-phase HPLC using a Phenomenex Luna 250 x 30 mm, 5 µm, 100 Å, C₁₈ (version 2) column maintained at 25 °C. Mobile phases were 0.2% formic acid in water (A) and 100 % methanol (B). The gradient was 35% B for 5 min followed by a ramp to 70% B over 40 min at 50 mL/min. Concentration of pooled fractions, followed by lyophilization, gave the **PA30-EDTA** conjugate (1.0 mg, 16% yield) as a white, fluffy solid: exact mass [M+H]⁺ = 2282.0657, experimental (ESI) [M+H]⁺ = 2282.1257. The EDTA conjugate of **PA31** was prepared *via* a virtually identical procedure to the one described for **PA30**, but with the following changes. **PA31** (4.0 mg) was substituted for **PA30**. The preparative HPLC purification parameters were identical to those described above except mobile phase A consisted of 0.2% trifluoroacetic acid in water. Concentration of pooled fractions, followed by

lyophilization gave the **PA31-EDTA** conjugate (2.3 mg, 55% yield) as an off-white fluffy solid: exact mass $[M+2H+Na]^+ = 2250.0007$, experimental (ESI) $[M+2H+Na]^+ = 2249.9562$. Analytical HPLC characterization was performed using a Phenomenex Jupiter Proteo (C12), 4.6 x 50 mm, 4 μ m, 90 Å maintained at 40°C. Mobile phases consisted of 0.1% formic acid in water (A) and 100 % acetonitrile (B). The gradient consisted of 5% B for 0.75 min followed by a ramp to 60% B over 3.25 min at 2 mL/min. Retention times were 2.641 min, 2.486 min and 2.437 min for **PA1**, **PA30** and **PA31** respectively.

PA1 Im-Py-Py- β -Py-Py-Py- γ -Py-Py- β -Py-Py-Py-Py- β -Ta (3 TFA):

^1H NMR (600 MHz, DMSO- d_6) δ = 10.46 (s, 1 H), 9.93 (s, 1 H), 9.91 (s, 4 H), 9.90 - 9.87 (m, 2 H), 9.86 (s, 1 H), 9.83 (s, 1 H), 9.63 (br. s., 2 H), 8.12 - 8.02 (m, 5 H), 7.98 - 7.92 (m, 1 H), 7.88 (br. s., 3 H), 7.40 (s, 1 H), 7.28 (d, J = 1.8 Hz, 1 H), 7.24 - 7.21 (m, 3 H), 7.21 - 7.18 (m, 4 H), 7.18 - 7.16 (m, 3 H), 7.15 (d, J = 1.8 Hz, 1 H), 7.08 (d, J = 1.8 Hz, 1 H), 7.07 (d, J = 1.8 Hz, 1 H), 7.06 (s, 1 H), 7.04 (d, J = 1.8 Hz, 1 H), 6.90 (d, J = 1.8 Hz, 1 H), 6.89 (d, J = 1.8 Hz, 1 H), 6.89 (s, 2 H), 6.85 (s, 2 H), 6.84 (d, J = 1.2 Hz, 1 H), 3.99 (s, 3 H), 3.85 (s, 3 H), 3.85 (s, 3 H), 3.84 (s, 9 H), 3.82 (s, 6 H), 3.81 (s, 3 H), 3.81 (s, 9 H), 3.49 - 3.42 (m, 4 H), 3.41 - 3.35 (m, 4 H), 3.25 - 3.16 (m, 4 H), 3.16 - 3.10 (m, 2 H), 3.10 - 2.98 (m, 4 H), 2.87 (br. s, 2 H), 2.74 (d, J = 4.7 Hz, 3 H), 2.55 - 2.51 (m, 2 H), 2.36 (t, J = 7.0 Hz, 2 H), 2.28 (t, J = 7.3 Hz, 2 H), 1.96 - 1.87 (m, 2 H), 1.82 - 1.74 (m, 2 H)

^{13}C NMR (151 MHz, DMSO- d_6) δ = 171.1, 169.3, 167.9, 167.8, 161.3, 158.6, 158.5, 158.5, 158.4, 158.4, 158.4, 158.1, 157.9, 155.9, 138.7, 126.8, 126.3, 123.0, 122.9, 122.8, 122.8, 122.8, 122.7, 122.7, 122.2, 122.2, 122.1, 122.1, 122.1, 122.0, 121.9, 121.4, 118.6, 118.5, 118.4, 118.2, 118.1, 118.0, 117.8, 104.9, 104.8, 104.8, 104.7, 104.3, 104.3, 104.2, 104.0, 104.0, 103.9, 53.3, 52.1, 40.0, 39.3, 38.2, 36.2, 36.1, 36.1, 36.0, 35.8, 35.8, 35.6, 35.6, 35.5, 35.4, 35.1, 33.3, 25.7, 24.0, 21.8

HRMS (ESI) calculated for $\text{C}_{91}\text{H}_{111}\text{N}_{31}\text{O}_{16}$ $[\text{MH}]^+$, 1894.8908, found, 1894.8933.

HPLC purity: 98%. (See **Figure SI 1a**).

PA30 TMG-Im-Py-Py- β -Py-Py-Py- γ -Py-Py- β -Py-Py-Py-Py- β -Ta (4 TFA):

^1H NMR (600 MHz, DMSO- d_6) δ = 10.19 (s, 1 H), 10.08 (s, 1 H), 9.94 (s, 1 H), 9.92 (s, 3 H), 9.91 (s, 1 H), 9.90 (s, 1 H), 9.87 (s, 1 H), 9.84 (s, 1 H), 9.45 (br. s., 1 H), 8.13 - 8.03 (m, 5 H), 7.86 (br. s., 1 H), 7.80 (br. s., 3 H), 7.29 (d, J = 1.8 Hz, 1 H), 7.23 (d, J = 1.2 Hz, 2 H), 7.22 - 7.21 (m, 2 H), 7.21 - 7.18 (m, 4 H), 7.16 (s, 3 H), 7.15 (d, J = 1.8 Hz, 1 H), 7.08 (d, J = 1.8 Hz, 1 H), 7.07 (d, J = 1.8 Hz, 1 H), 7.05 (d, J = 1.2 Hz, 1 H), 6.91 (d, J = 1.8 Hz, 1 H), 6.89 (d, J = 1.8 Hz, 2 H), 6.88 (s, 1 H), 6.85 (s, 2 H), 6.84 (d, J = 1.2 Hz, 1 H), 6.56 (br. s., 2 H), 3.99 (s, 3 H), 3.85 (s, 3 H), 3.85 (s, 3 H), 3.84 (s, 3 H), 3.84 (s, 3 H), 3.84 (s, 6 H), 3.81 (s, 6 H), 3.81 (s, 3 H), 3.80 (s, 6 H), 3.50 - 3.26 (m, 6 H), 3.24 - 3.16 (m, 3 H), 3.16 - 3.09 (m, 2 H), 3.09 - 2.98 (m, 3 H), 2.98 - 2.80 (m, 14 H), 2.74 (d, J = 5.3 Hz, 3 H), 2.55 - 2.51 (m, 4 H), 2.36 (t, J = 7.3 Hz, 2 H), 2.27 (t, J = 7.3 Hz, 2 H), 1.94 - 1.85 (m, 2 H), 1.82 - 1.73 (m, 4 H)

^{13}C NMR (151 MHz, DMSO- d_6) δ = 171.0, 169.3, 167.9, 167.8, 161.3, 158.5, 158.5, 158.4, 158.4, 158.4, 158.2, 157.9, 157.8, 155.4, 135.9, 134.8, 123.1, 122.9, 122.8, 122.8, 122.8, 122.7, 122.7, 122.2, 122.2, 122.1, 122.1, 122.1, 122.1, 122.0, 121.9, 121.1, 118.9, 118.5, 118.5, 118.2, 118.1, 118.0, 117.8, 115.8, 105.0, 104.8, 104.8, 104.7, 104.3, 104.3, 104.0, 104.0, 103.9, 53.3, 52.1, 40.0, 39.3, 38.2, 36.2, 36.1, 36.1, 36.0, 35.8, 35.8, 35.6, 35.6, 35.5, 35.4, 33.3, 25.7, 24.0, 22.5, 21.8

HRMS (ESI) calculated for $C_{96}H_{122}N_{34}O_{16} [M]^+$, 2006.9778, found, 2006.9653.

HPLC purity: 99%. (See **Figure SI 1b**).

PA31 Guan-Im-Py-Py-β-Py-Py-Py-γ-Py-Py-β-Py-Py-Py-Py-β-Ta (4 TFA):

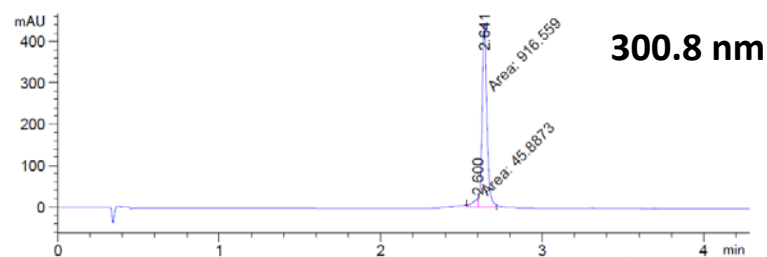
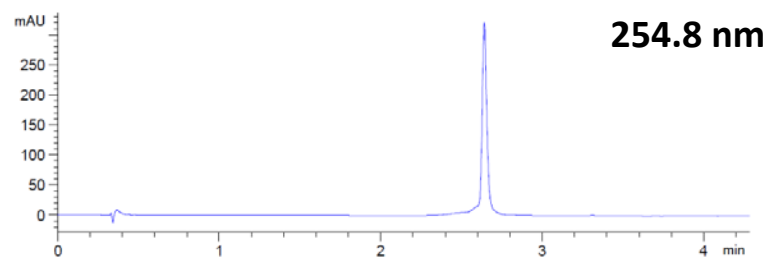
1H NMR (600 MHz, DMSO- d_6) δ = 10.46 (s, 1 H), 10.42 (s, 1 H), 9.94 (s, 1 H), 9.93 (s, 1 H), 9.92 (s, 3 H), 9.91 (s, 1 H), 9.90 (s, 1 H), 9.87 (s, 1 H), 9.84 (s, 1 H), 9.46 (br. s., 1 H), 8.13 - 8.07 (m, 4 H), 8.06 (t, J = 5.6 Hz, 2 H), 7.86 (br. s., 1 H), 7.81 (br. s., 3 H), 7.28 (d, J = 1.8 Hz, 1 H), 7.23 (d, J = 1.2 Hz, 2 H), 7.22 (d, J = 1.2 Hz, 1 H), 7.21 - 7.18 (m, 5 H), 7.17 (s, 3 H), 7.14 (d, J = 1.8 Hz, 1 H), 7.08 (d, J = 1.2 Hz, 1 H), 7.07 (d, J = 1.8 Hz, 1 H), 7.04 (d, J = 1.8 Hz, 1 H), 6.91 (d, J = 1.8 Hz, 1 H), 6.89 (s, 2 H), 6.88 (d, J = 1.8 Hz, 1 H), 6.87 - 6.85 (m, 2 H), 6.84 (d, J = 1.8 Hz, 1 H), 6.69 (s, 1 H), 6.56 (br. s., 2 H), 4.00 (s, 3 H), 3.86 (s, 3 H), 3.85 (s, 3 H), 3.84 (s, 3 H), 3.84 (s, 3 H), 3.84 (s, 6 H), 3.81 (s, 6 H), 3.81 (s, 3 H), 3.80 (s, 6 H), 3.51 - 3.32 (m, 6 H), 3.24 - 3.16 (m, 3 H), 3.16 - 3.09 (m, J = 6.2, 12.5, 12.5 Hz, 2 H), 3.09 - 2.98 (m, 3 H), 2.90 - 2.81 (m, 2 H), 2.74 (d, J = 4.7 Hz, 3 H), 2.55 - 2.51 (m, 4 H), 2.36 (t, J = 7.3 Hz, 2 H), 2.27 (t, J = 7.3 Hz, 2 H), 1.94 - 1.85 (m, 2 H), 1.82 - 1.75 (m, 4 H)

^{13}C NMR (151 MHz, DMSO- d_6) δ = 171.1, 169.3, 167.9, 161.3, 161.3, 158.6, 158.5, 158.5, 158.4, 158.4, 158.3, 158.2, 158.0, 155.2, 154.6, 134.7, 134.4, 123.1, 122.9, 122.8, 122.8, 122.7, 122.7, 122.2, 122.2, 122.1, 122.1, 122.1, 122.0, 121.9, 120.9, 118.9, 118.5, 118.4, 118.2, 118.1, 118.0, 118.0, 117.8, 116.1, 113.7, 105.1, 104.8, 104.8, 104.7, 104.3, 104.3, 104.2, 104.0, 104.0, 103.9, 53.3, 52.1, 40.0, 39.3, 38.2, 36.2, 36.2, 36.1, 36.0, 36.0, 35.8, 35.8, 35.6, 35.6, 35.5, 35.4, 35.4, 33.3, 25.7, 24.1, 22.5, 21.8

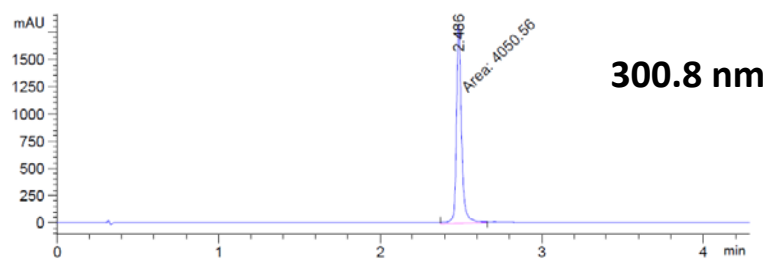
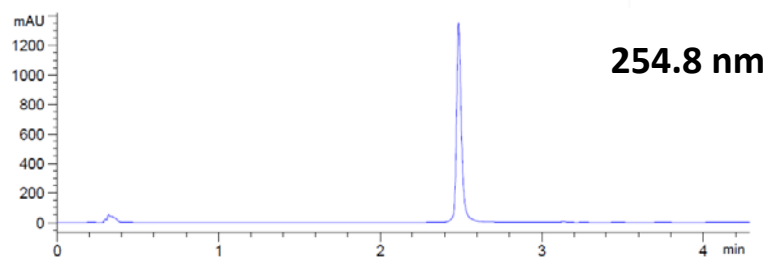
HRMS (ESI) calculated for $C_{92}H_{114}N_{34}O_{16} [M]^+$, 1950.9152, found, 1950.9034.

HPLC purity: 98%. (See **Figure SI 1c**).

a PA1



b PA30



c PA31

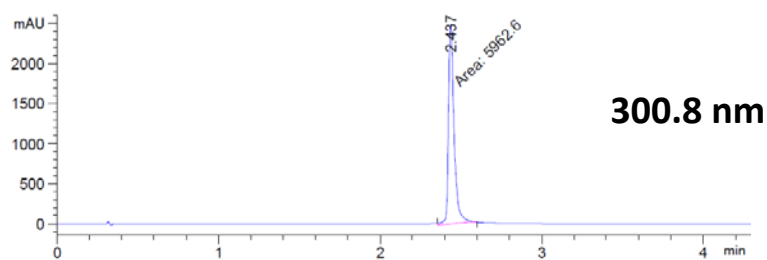
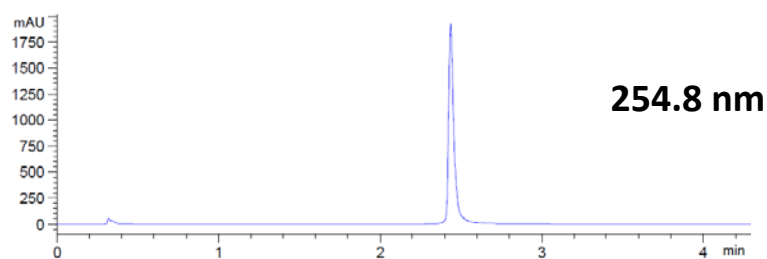


Figure SI 1 a-c. HPLC Purity of PA1, PA30 and PA31 with λ monitored.

II. Compound efficacy testing

The antiviral activities were calculated by measuring the ability of PAs to decrease the viral DNA (or episomal) load in monolayer keratinocyte cultures. The episomal load was determined by Q-PCR as previously described.^[9] **PA1** was already reported, along with extensive discussion of how antiviral parameters are defined for this new mechanism of antiviral action in which the viral genome is degraded while the host cells are unharmed as measured by standard MTT and LDH assays and observations of numbers of floating vs. adherent cells, etc.^[8,9,10] Typical assays were run for 48 h though much longer assays have also been run, and ref. 9 also describes organotypic, differentiated keratinocyte tissue culture assays as well as cell monolayer assays. Because **PA1** has been central to our work for over a decade, it has been used as an internal standard to benchmark new compound series, so even though we cite the prior study, **PA1** was studied side-by-side with **PA30** and **PA31** for the current work. The new **PA1** results were indistinguishable from the dose-response curves previously reported. Dose-response curves for **PA30** and **PA31** are given below in **Figure SI 2a** and **2b**.

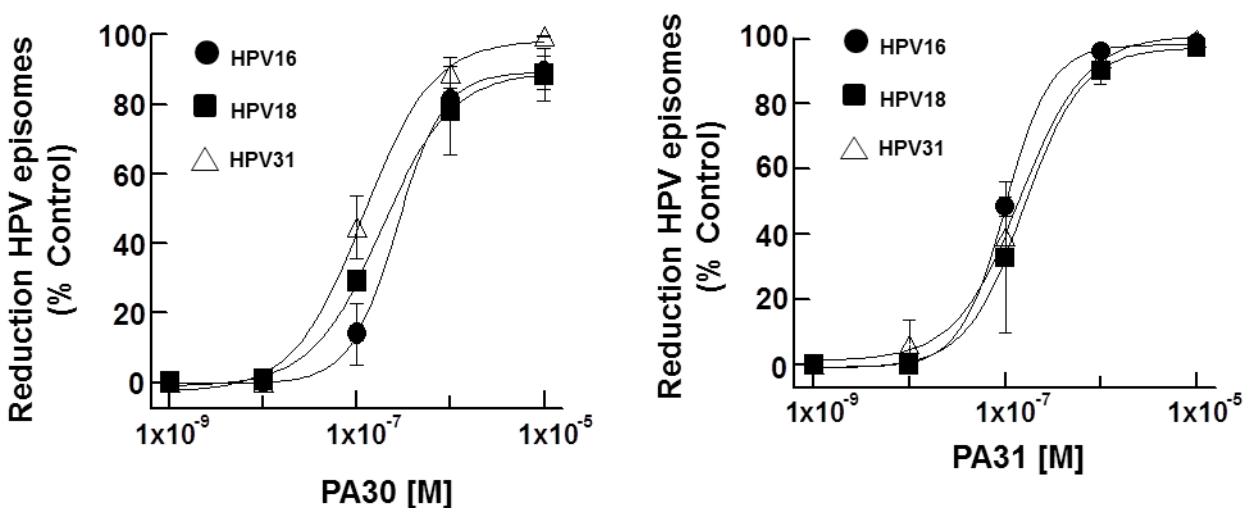
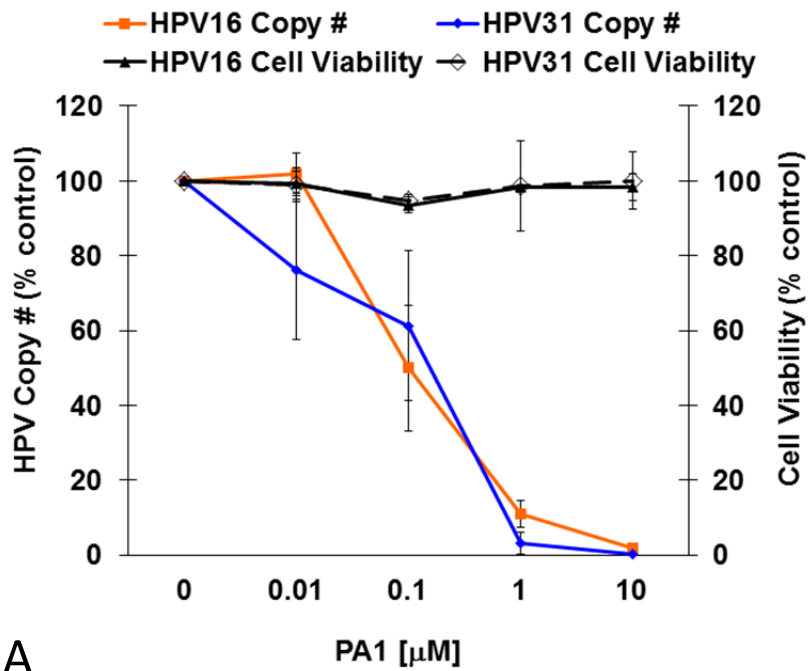
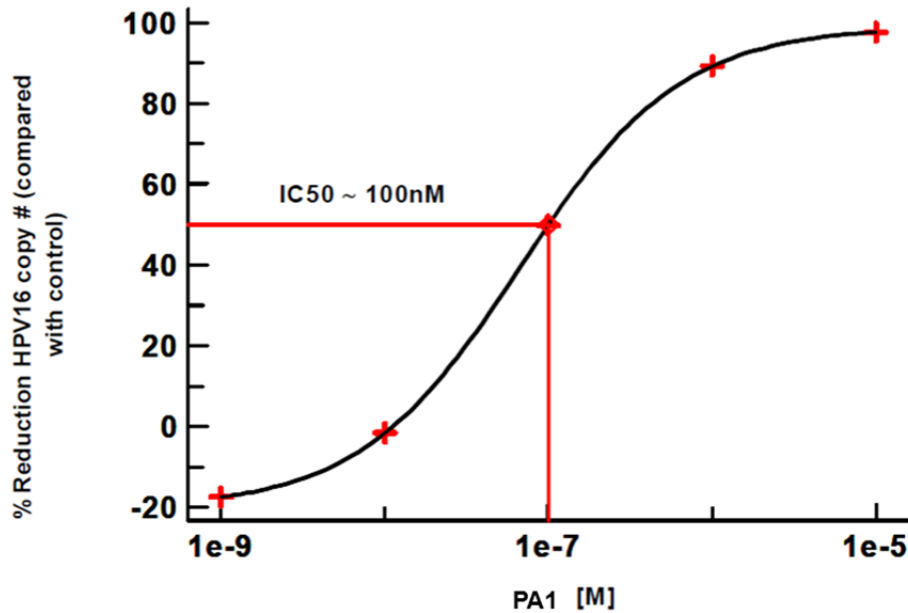


Figure SI 2. Dose-response curves for PA30 and PA31 against HPV16, 18 and 31 from which IC_{50} and IC_{90} values were determined as polyamide concentrations where 50 and 90% of viral DNA was undetectable relative to vehicle control. (a) Results for PA30. (b) Results for PA31.

Note that the plots show reduction, or decrease, in HPV DNA as a % of control, so 100% reduction corresponds to elimination of viral DNA on the scale plotted. Vehicle control was 0.1% DMSO in water, which is well tolerated by the keratinocytes.^[8,9,10] Differences in efficacy between **PA30** and **PA31** are striking when HPV16 and 18 are considered: **PA30** (**Figure SI 2a**) achieves close to 100% reduction in viral DNA only for HPV31, a high-risk but minor form of the disease that is still clinically important. **PA31** is highly-active and essentially indistinguishably-active against all three high-risk HPVs, on the scale plotted in **Figure SI 2b**. Plots for **PA1** are shown in **Figure SI 3** below.



A



B

Q-PCR; n = 4 independent cell culture experiments in the presence of increasing concentrations of PA1.

Figure SI 3. Dose-response curves for **PA1** against HPV16 and 31 from which IC_{50} and IC_{90} values were determined. (A) Results for HPV16 and 31 with cell toxicity data. (b) Expanded detail of results for HPV16, plotted as the inverse of the other results (as the percent decrease of viral DNA vs. control, reaching 100% decrease at maximum efficacy, instead of the simple decrease in DNA vs. control).

III. Quantitative DNase I footprinting

The molar extinction coefficient for **PA1** was determined as $\epsilon = 88,235 \text{ M}^{-1}\text{cm}^{-1}$ by fluorescence as previously described.^[53] The molar extinction coefficients for **PA30** and **PA31** were determined as $\epsilon = 91,164 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon = 89,263 \text{ M}^{-1}\text{cm}^{-1}$ by measuring the absorbance at 305 nm of solutions with known polyamide concentrations in the range $5 \times 10^{-7} \text{ M}$ to $7 \times 10^{-5} \text{ M}$.

Quantitative DNase I footprinting experiments were performed as described previously.^[18] Briefly, a 305 bp DNA fragment corresponding to the nucleotides 7479-7783 in the Long Control Region (LCR) of the HPV18 genome (Genbank accession number X05015) was PCR amplified and fluorescently-labeled at both 5' ends using the following primers: (top strand) 5'-FAM-CT TAT GTC TGT GGT TTT CTG and (bottom strand) 5'-HEX-TT CAT GTT AAG GGT AGA CAG. The DNA fragment was diluted to 200 pM in TKMC buffer with 10 mM CHAPS and incubated with different concentrations of each polyamide (0.5-40 nM) for at least 4 h at 37 °C. The polyamide solutions were always added to DNA solutions not *vice versa*. At the end of this incubation period, each solution was subjected to ~0.08 U of DNase I for 5 min. The reactions were then quenched with EDTA and the DNA fragments were purified using a QIAquick PCR purification kit. Each experiment was performed at least three times.

Table SI 1. K_d Values Determined by Hill Equation for Sites 1-3 with Hill Coefficients and Overall Fit

	PA1			PA30			PA31		
	K_d (Hill)	Hill Coeff.	R	K_d (Hill)	Hill Coeff.	R	K_d (Hill)	Hill Coeff.	R
Site 1	0.9 ± 0.1	1.6 ± 0.4	>0.99	1.2 ± 0.2	1.8 ± 0.2	>0.99	1.0 ± 0.1	2.0 ± 0.3	>0.99
Site 2	0.9 ± 0.2	1.8 ± 0.6	1	1.1 ± 0.2	2.0 ± 0.3	>0.98	1.0 ± 0.1	1.6 ± 0.2	>0.99
Site 3	2.7 ± 0.5	3.8 ± 0.8	>0.99	3.2 ± 0.3	3.4 ± 0.5	>0.98	2.2 ± 0.2	3.8 ± 0.8	1

Table SI 2. K_d Values Determined by Langmuir Equation for Sites 1-3 with Overall Fit

	PA1		PA30		PA31	
	K_d (Langmuir)	R	K_d (Langmuir)	R	K_d (Langmuir)	R
Site 1	1.0 ± 0.1	>0.98	1.4 ± 0.3	>0.98	1.3 ± 0.4	>0.97
Site 2	1.0 ± 0.2	>0.97	1.3 ± 0.3	>0.96	1.3 ± 0.3	>0.99
Site 3	3.5 ± 0.5	>0.92	6 ± 1	>0.95	2.9 ± 1.5	>0.95

IV. Affinity cleavage

Affinity cleavage experiments were performed as previously described. Briefly, PA-EDTA conjugates were complexed with 0.8 eq of Fe^{2+} in the form of ammonium iron(II) sulfate hexahydrate. 1 nM DNA fragment in 10 mM Tris, 10 mM CHAPS, pH 7.5 was incubated with different concentrations of

PA-EDTA-conjugates (5-200 nM) for at least 4 h at 37 °C. The reaction was initiated with 5 µL of 100 mM DTT and incubated at room temperature for 1-2 h. Each cleavage reaction was then quenched and fragments were purified using a QIAquick PCR purification kit.

The samples from quantitative DNase I footprinting and affinity cleavage experiments were analyzed using capillary electrophoresis and indexed using Sanger and Maxam-Gilbert sequencing, respectively.^[SI4] Dissociation constant (K_d) values were calculated by normalizing the area from a peak in the footprint region to the area of a peak not sensitive to PA concentration. These values were plotted and fitted to Hill and Langmuir binding isotherms (as appropriate) with KaleidaGraph in methods reported in detail elsewhere.^[18, SI4 and SI5]

Additional references cited in SI.

- SI1.** M. del Fresno, A. El-Faham, L. A. Carpino, M. Royo, F. Albericio, *Org. Lett.*, 2000, **2**, 3539.
- SI2.** S. Robinson, E. J. Roskamp, *Tetrahedron*, 1997, **53**, 6697.
- SI3.** C. M. Dupureur, J. K. Bashkin, K. Aston, K. J. Koeller, K. R. Gaston, G. He, *Anal. Biochem.*, 2012, **423**, 178.
- SI4.** G. He, E. Vasilieva, J. K. Bashkin, C. M. Dupureur, *Anal. Biochem.*, 2013, **439**, 99.
- SI5.** K. J. Koeller, G. D. Harris, J., K. Aston, G. He, C. H. Castaneda, M. A. Thornton, T. G. Edwards, S. Wang, R. Nanjunda, W. D. Wilson, W. D., C. Fisher, J. K. Bashkin, *Med. Chem.*, 2014, **4**, 338.