Electronic Supplementary Material (ESI) for Chemical Communications.

Supplementary Information

Exploiting anthracene photodimerization within peptides: light induced sequence-selective DNA binding

Gemma A. Bullen, James H. R. Tucker and Anna F. A. Peacock*

School of Chemistry, University of Birmingham, Edgbaston, Birmingham, B15 2TT

*To whom correspondence should be addressed, email: a.f.a.peacock@bham.ac.uk

Contents:

Acknowledgements, Methods and Materials, one supplementary scheme, two supplementary tables and eight supplementary figures.

1. Acknowledgements

We thank the School of Chemistry, the University of Birmingham and the ESPRC for a Ph.D. studentship (G.A.B.) and support of this research (Leadership fellowship to JHRT, EP/G007578/1, Functional DNA-based assemblies). Thanks also to M. R. Berwick for development of schematic images. Some equipment used in this research was obtained, through Birmingham Science City: Innovative Uses for Advanced Materials in the Modern World (West Midlands Centre for Advanced Materials Project 2), with support from Advantage West Midlands (AWM) and part funded by the European Regional Development Fund (ERDF).

2. Methods and Materials

Mono- and di hydrogen potassium salts, sodium chloride, acrylamide, tetramethylethylenediamine, ammonium persulphate, Tris base, hydrochloric acid 32%, sodium dodecyl sulphate, ethylene diamine tetra acetic acid, magnesium chloride, potassium chloride, trifluoroacetic acid, glacial acetic acid and triethyl amine were all purchased from Fisher Scientific Ltd. Tricine was obtained from Alfa Aesar; anisole, thioanisole and 1,2-ethanedithiol were purchased from Merek Millipore. Sigma Aldrich Ltd. were used to supply glycerol, Igpeal CA 630, bovine serum albumin and 9-anthracenecarboxylic acid. SybrGold was sourced from Invitrogen and coomassie blue stain from AppliChem. Fmoc-Arg(Pbf)-OH, Fmoc-Leu-OH, Fmoc-Met-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, HBTU, DIPEA and rink amide MBHA resin were purchased from AGTC Bioproducts Ltd. Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Lys(Boc)-OH, 20% piperidine in DMF and synthesis grade DMF were purchased from Pepceuticals Ltd. Fmoc-Lys(Mtt)-OH was obtained from Nova Biochem Ltd. All oligonucleotide synthesis reagents were purchased from Link Technologies Ltd. Unless otherwise stated, all reagents were used as per the manufacturers instructions without further purification. Errors are reported as standard deviations. MilliQ water was used for all studies.

Peptide synthesis and purification

Ac-<u>ALKRARNTEAARRSR</u>ARKLQRMKQ**<u>KG</u>**-NH_2

The peptide was synthesized using standard solid-phase peptide synthesis (SPPS) methods[1] on rink amide MBHA resin (0.1 mM scale, 0.65 mmol/g loading) using a CEM Liberty 1 automated peptide synthesizer. Initial couplings (bold and underlined) were performed manually to allow for the retrieval of Fmoc-Lys(Mtt)-OH. The 4-methyltrityl (Mtt) protecting group was removed with 3% TFA in DCM for 10 minutes with 3 repeats. The resin was then rinsed 5 times with 10 mL DCM and 5 times with 10 mL DMF. A tenfold excess of 9-anthracenecarboxylic acid was then coupled using SPPS synthesis methods, a 15 hour coupling was required. The resin was then loaded into the CEM Liberty 1 automated peptide synthesizer and the remaining amino acids coupled. A standard coupling programme (25 W, 75 ºC, 300 s) was used for all amino acids except for the coupling of Fmoc-

Arg(Pbf)-OH and Fmoc-Ser(tBu)-OH which were double coupled. Residues towards the N-terminus (underlined) were double coupled. Deprotections were performed in 20% piperidine in DMF (35 W, 75 ºC, 180 s). The N-terminus was manually acetylated (20% acetic anhydride, 20% DIPEA in DMF) before cleavage from the resin and deprotection of side chains (90% trifluoroacetic acid, 5% thioanisole, 3% 1, 2-ethanedithiol and 2% anisole for 2 hours). The peptide was purified and characterised as previously described.[2] Peptide concentrations were calculated based on the anthracene absorbance at 375 nm, $\varepsilon_{375\ nm}$ = 6300 M^{-1} cm^{-1} in water. Subsequent studies were performed using 2 μM anthracene tagged peptide in 100 mM sodium chloride and 10 mM sodium phosphate buffer pH 7.

Oligonucleotide Synthesis

NS DNA 1	5'-TGGAGTATGCGTCGATTCGT-3'
NS DNA 2	3'-ACCTCATACGCAGCTAAGCA-5'
CRE DNA 1	5'-ACGAGATGACGTCATCTCCA-3'
CRE DNA 2	3'-TGCTCTACTGCAGTAGAGGT-5'
Half CRE DNA 1	5'-TGGAGATGACGTTGTCTCGT-3'
Half CRE DNA 2	3'-ACCTCTACTGCAACAGAGCA-5'

The strands were synthesized using standard oligonucleotide synthesis protocols using an Applied Biosciences ABI 395 synthesizer. Oligonucleotides were purified and characterized as previously described.[3]

Irradiation studies

Samples were irradiated with a water cooled 125 W mercury arc lamp, purchased from Photochemical Reactors Ltd, placed in a housing with a small outlet fitted with a 365 nm filter purchased from Edmunds Optics. All irradiations were carried out for 10 minutes.

Optical studies

Fluorescence studies were performed on a Shimadzu RF-5301 PC spectrofluorimeter. Samples were excited at 325 nm with a slit width of 3 nm and an emission slit width of 5 nm. Absorbance studies were performed on a Shimadzu UV-1800 spectrophotometer. All studies were performed in a low volume 1 cm pathlength quartz emission cell.

Circular dichroism studies

Circular dichroism studies were performed on a Jasco J-810 spectropolarimeter using a 1 cm pathlength quartz cuvette. A data pitch of 0.2 nm was used at a scan rate of 200 nm/min with a response time of 1 second, with 10 accumulations per sample. Molar ellipticity values were calculated using Equation 1;

$$\Theta_m = \frac{\Theta}{10 \times l \times c \times N_a} \quad (1)$$

where Θ_m is the molar ellipticity, Θ is the ellipticity, l is the pathlength in cm, c is the concentration of anthracene tagged peptide in mol dm^{-3} and N_a is the number of amino acid residues in the peptide. Percentage folded values were calculated based on the molar ellipticity at 222 nm using Equation 2;[4]

$$\% \ folding = \left(\frac{\Theta_m - 640}{\left(-42500 \times \left(1 - \left(\frac{3}{N_a} \right) \right) \right) - 640} \right) \times 100 \quad (2)$$

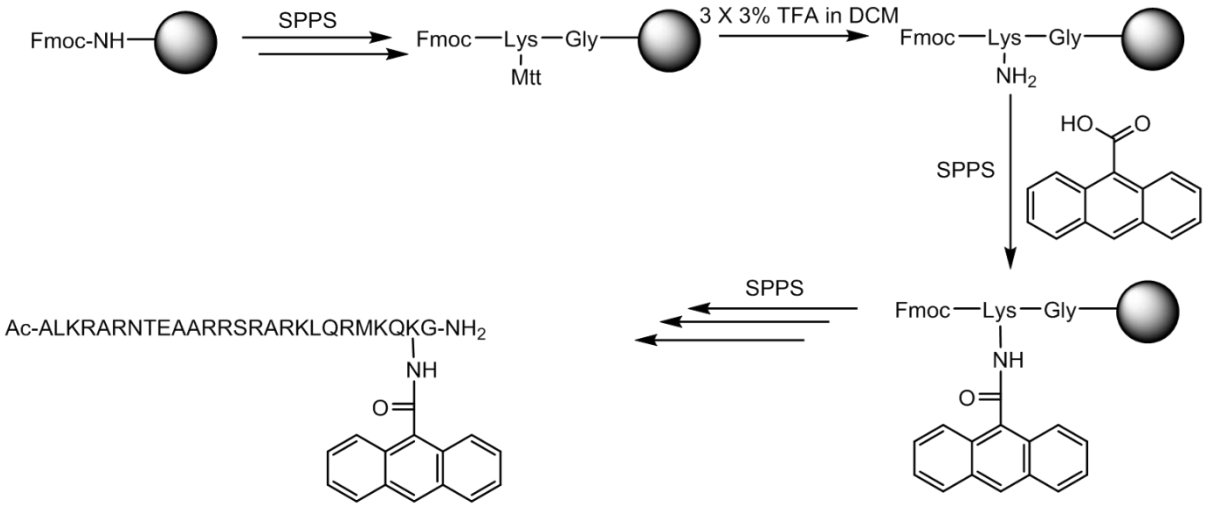
Denaturing Gel Electrophoresis

SDS PAGE gels were performed with a 20% polyacrylamide separating gel with 0.5 M Tris HCl buffer pH 8.8, 0.5% SDS and a 4% polyacrylamide stacking gel with 0.1 M Tris HCl pH 6.8, 2% SDS. Samples were prepared in 250 mM Tris HCl pH 6.8, 10% glycerol and 2% SDS and equilibrated for 1 hour. Gel was run with 1x Tris Tricine running buffer and stained with coomassie brilliant blue R-250 protein stain.

Gel Electrophoretic Mobility Shift Assays (EMSA)

Gel EMSA studies were performed using a 10% polyacrylamide gel with 0.5x TBE buffer. 0.5x TBE buffer was also used as a running buffer. Peptide and DNA were prepared in 20 mM Tris HCl pH 7.5, 2 mM $MgCl_2$, 2 mM EDTA, 90 mM KCl, 10 % glycerol, 2.25 % Igpeal CA630 and 0.1 mg/mL BSA. Samples were then irradiated with 365 nm light for 10 minutes where stated, and equilibrated for 1 hour. Gels were visualised with Sybr Gold stain.

3. Scheme S1



Scheme S1. Peptide synthesis schematic showing the manual couplings of the first glycine and Mtt protected lysine followed by 9-anthracenecarboxylic acid. The resin was then loaded into an automated synthesizer for the rest of the sequence to be synthesized.

4. Figure S1

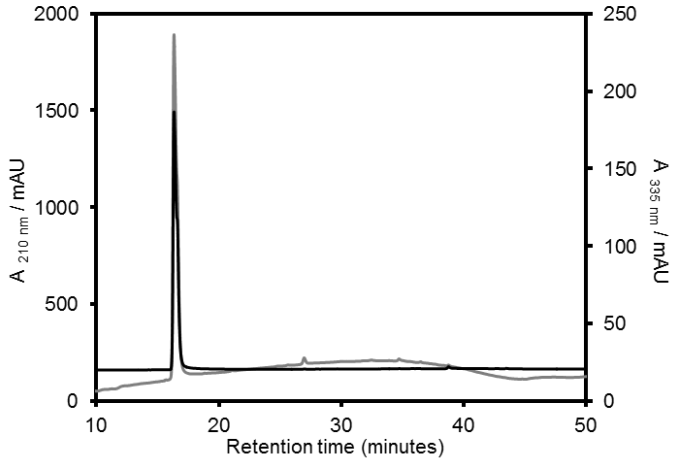


Figure S1. Analytical HPLC of peptide, monitored at 210 nm (black) for peptide and at 335 nm (grey) for anthracene. Run with a gradient of 0-100% MeCN over 40 minutes with 0.1% TFA followed by a 10 minute hold with 100% MeCN with 0.1% TFA before returning to 0% MeCN and holding for 10 minutes.

5. Figure S2

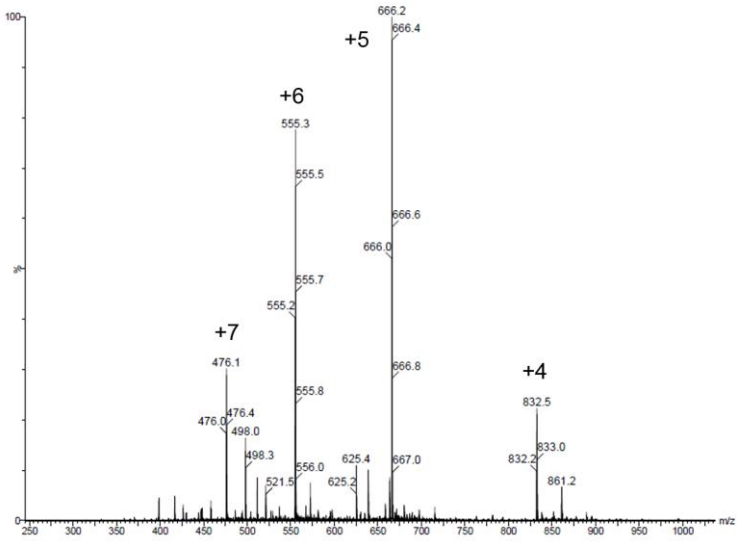


Figure S2. Positive mode electrospray mass spectrometry of the purified anthracene tagged peptide showing the charge envelope (MW 3326).

6. Table S1

Table S1. DNA sequences used in this study with the CRE and half CRE sites shown in bold with the expected and observed masses from ESI mass spectrometry.

DNA Stand	Sequence	Expected Mass	Observed Mass
CRE 1	5'-ACGAG**ATGACGTCAT**CTCCA-3'	6086	6087
CRE 2	3'-TGCTC**TACTGCAGTA**GAGGT-5'	6148	6148
Half CRE 1	5'-TGGAG**ATGACG**TTGTCTCGT-3'	6179	6179
Half CRE 2	3'-ACCTC**TACTGC**AACAGAGCA-5'	6055	6055
NS 1	5'-TGGAGTATGCGTCGATTCGT-3'	6179	6179
NS 2	3'-ACCTCATACGCAGCTAAGCA-5'	6055	6057

7. Figure S3

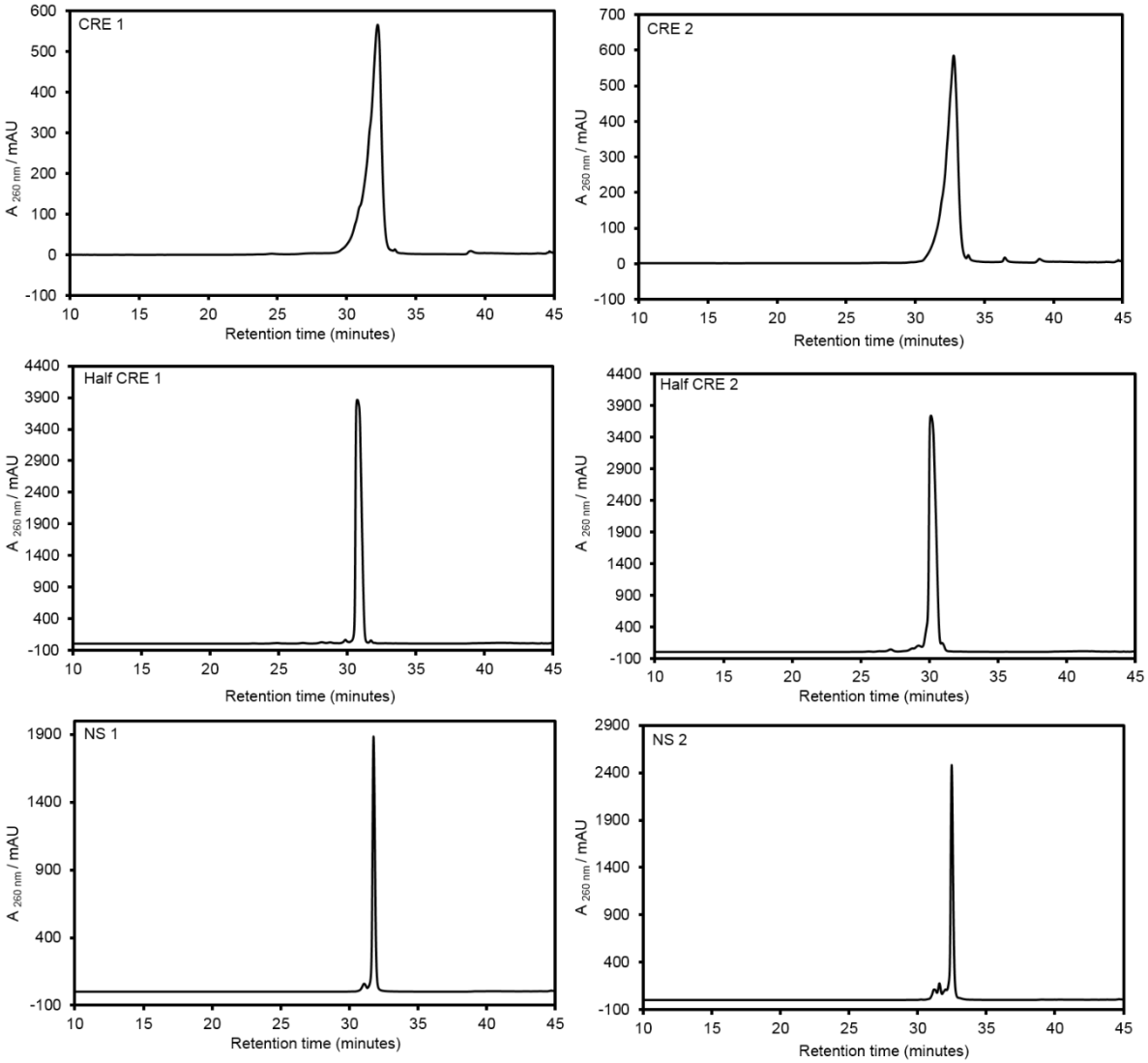


Figure S3. Analytical HPLC of DNA strands monitored at 260 nm. Run with linear gradients of 20 -> 50% solvent B (0-25 mins); 50 -> 100% solvent B (25-35 mins); 0 -> 100% solvent C (35-45 mins); hold with 100% solvent C (45-55 mins); return to 30% solvent B and hold (55-60 mins). Where solvent A (5% MeCN in 0.1 mM TEAA) made the remaining solvent volumes, solvent B is 15% MeCN in 0.1 mM TEAA and solvent C is 100% MeCN.

8. Figure S4

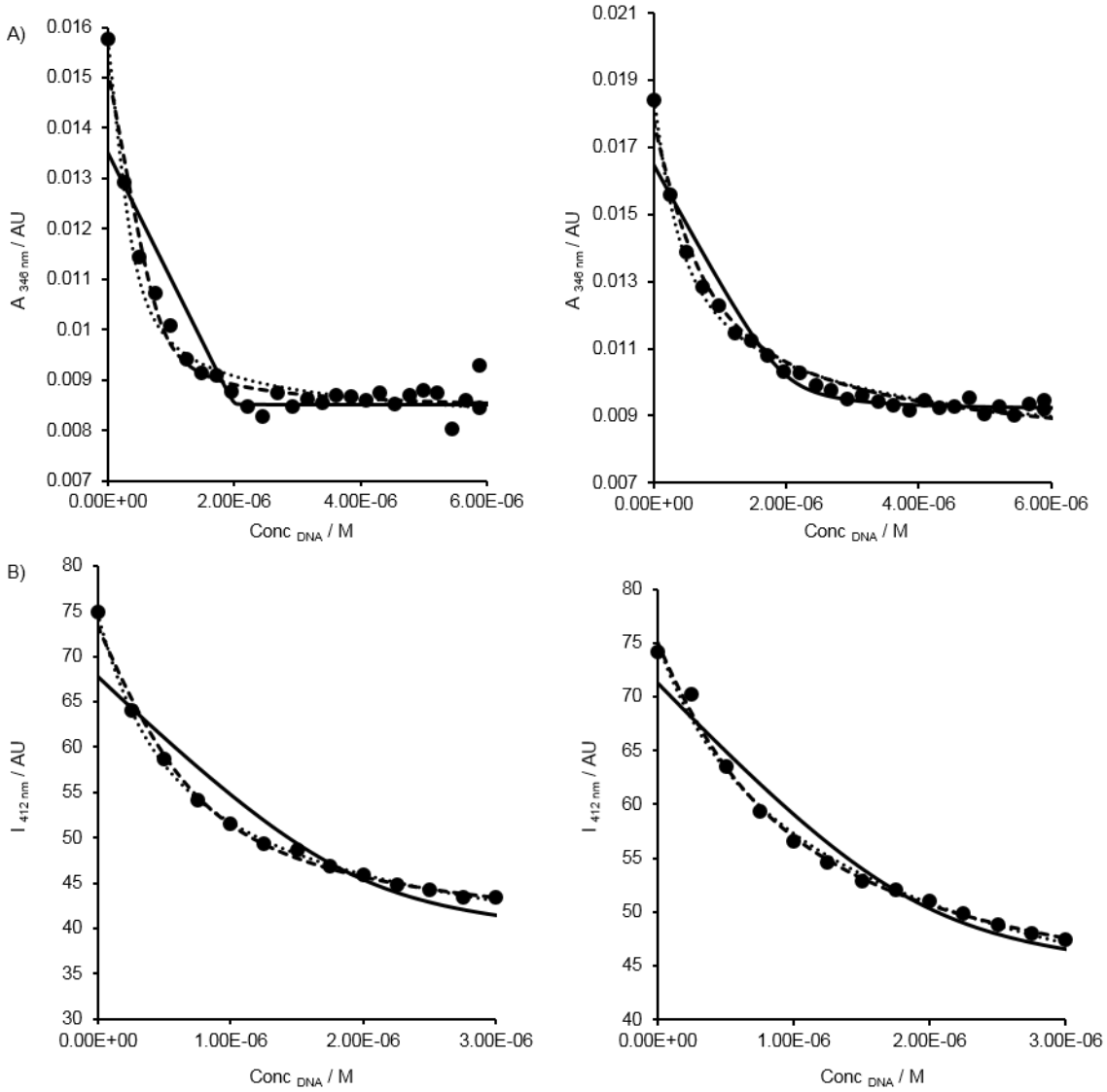


Figure S4. Fittings for titrations of DNA into a 2 μM anthracene tagged peptide solution, monitored by (A) UV-Vis and (B) fluorescence. Titrations of CRE DNA (left) fit to a 1:1 (solid), 2:1 (dash, UV-vis logK_D -6.7; fluorescence logK_D -6.1) and 3:1 (dotted, UV-vis logK_D -6.2; fluorescence logK_D -5.8) peptide:DNA binding models. Titrations of NS DNA (right) fit to a 1:1 (solid), 2:1 (dash, UV-vis logK_D -6.0; fluorescence logK_D -5.9) and 3:1 (dotted, UV-vis logK_D -5.9; fluorescence logK_D -5.6) peptide:DNA binding models. All fittings performed using Dynafit software (Biokin Ltd., Massachusetts, USA).[5] All titrations performed in 10 mM sodium phosphate buffer pH 7 and 100 mM NaCl at 298 K.

9. Figure S5

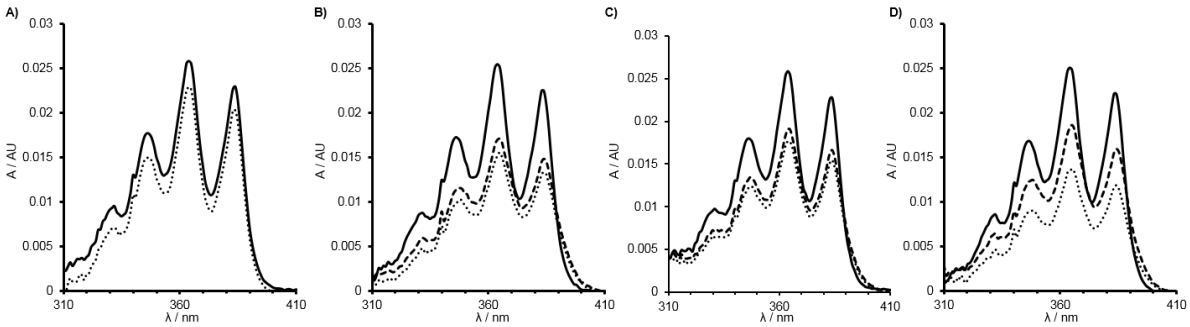


Figure S5. Absorbance spectra of 2 μM solutions of anthracene tagged peptide (solid), with 1 μM duplex DNA, where applicable, (dash) and after 10 minute irradiation with 365 nm light (dot) in the absence (A) and the presence of NS (B), half CRE (C) and CRE (D) DNA, in 10 mM sodium phosphate buffer pH 7 and 100 mM NaCl at 298 K.

10. Figure S6

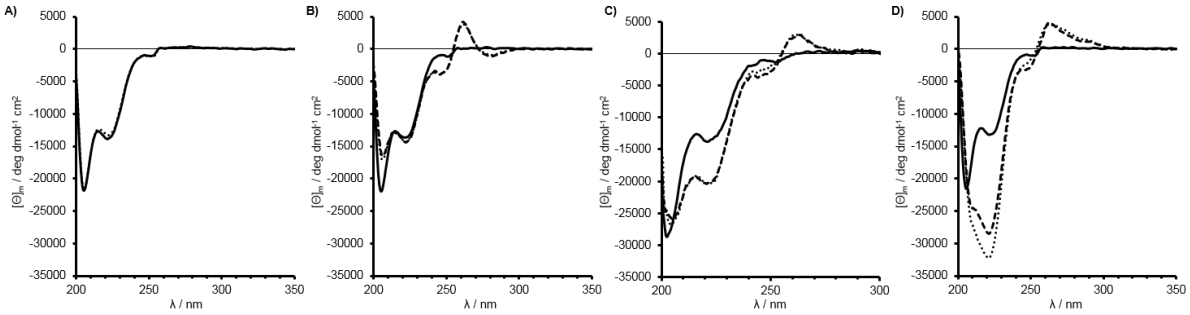


Figure S6. Induced CD spectra of 2 μM solutions of anthracene tagged peptide (solid), with 1 μM duplex DNA, where applicable, (dash) and after 10 minute irradiation with 365 nm light (dot) in the absence (A) and the presence of NS (B), half CRE (C) and CRE (D) DNA, in 10 mM sodium phosphate buffer pH 7 and 100 mM NaCl at 298 K.

11. Table S2

Table S2. Percentages changes in relative absorbance at 364 nm, fluorescence at 426 nm and folding from measure of ellipticity at 222 nm after irradiation with 365 nm light for 10 minutes with error calculated as standard deviation. Initial values for DNA containing samples were taken after the addition of DNA.

Sample	% change in absorbance at 364 nm (%)	% change in fluorescence at 426 nm (%)	% folded before irradiation (%)	% folded after irradiation (%)
No DNA	-8 ± 3	-7 ± 2	38 ± 1	36 ± 1
NS DNA	-9 ± 2	-8 ± 3	40 ± 2	40 ± 1
Half CRE DNA	-6 ± 2	-4 ± 3	57 ± 2	56 ± 2
CRE DNA	-27 ± 1	-20 ± 1	75 ± 1	84 ± 1

12. Figure S7

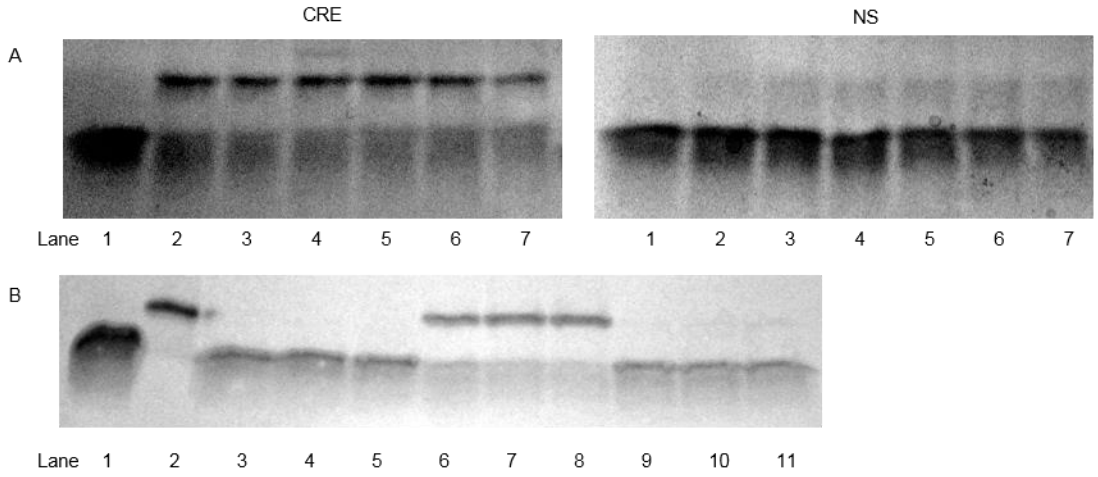


Figure S7. SDS PAGE tracking photodimer formation of anthracene tagged peptide (2 μM) over time with 1 μM CRE and NS DNA where applicable. A - Samples were run 90 min after sample preparation, having been irradiated with 365 nm light for the following times: Lane 1 – 0 mins, 2 – 15 mins, 3 – 30 mins, 4 – 45 mins, 5 – 60 mins, 6 – 75 mins and 7 – 90 mins. B – Samples run 30 min after sample preparation lane 1 – anthracene tagged peptide, 2 – $(GCN4bd1)_2Pyr$ covalent dimer control, the following lanes run from aliquots taken after 10, 20 and 30 minutes irradiation, 3-5 – peptide alone, 6-8 – peptide with CRE DNA and 9-11 – peptide with NS DNA. Visualized using coomassie brilliant blue R-250 protein stain.

13. Figure S8

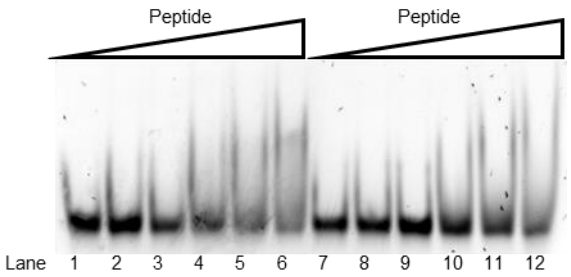


Figure S8. Gel EMSA assay of DNA recognition. Peptide irradiated for 10 minutes with 365 nm light prior to the addition of DNA then equilibrated for 1 hour. Lanes 1-6: CRE DNA (100 nM) and lanes 7-12 NS DNA (100 nM) with 0, 100, 200, 400, 600 and 1000 nM anthracene tagged peptide. DNA visualized using SybrGold.

14. Supporting Information References

[1] P. D. White, W. C.; Chan, *Fmoc Solid Phase Peptide Synthesis*, Oxford University Press: New York, New York, 2000.

[2] A. F. A. Peacock, G. A. Bullen, L. A. Gethings, J. P. Williams, F. H. Kriel, J. Coates, *J. Inorg. Biochem.* 2012, **117**, 298–305.

[3] J.-L. H. A. Duprey, Z. Zhao, D. M. Bassani, J. Manchester, J. S. Vyle, J. H. R. Tucker, *Chem. Commun.* 2011, ***47,*** 6629–31.

[4] J. K. Myers, C. N. Pace, J. M. Scholtz, *Proc. Natl. Acad. Sci.* 1997, **94**, 2833–7.

[5] P. Kuzmic, *Anal. Biochem.* 1996, **237**, 260-273.