

## Supplementary information

### Structural insights into the EthR-DNA interaction from native mass spectrometry

Daniel Shiu-Hin Chan,<sup>a</sup> Wei-Guang Seetoh,<sup>a</sup> Brendan N. McConnell,<sup>a</sup> Dijana Matak-Vinković,<sup>a</sup> Sherine E. Thomas,<sup>b</sup> Vitor Mendes,<sup>b</sup> Michal Blaszczyk,<sup>b</sup> Anthony G. Coyne,<sup>a</sup> Tom L. Blundell,<sup>b</sup> Chris Abell\*<sup>a</sup>

<sup>a</sup>Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW (United Kingdom)

E-mail: [ca26@cam.ac.uk](mailto:ca26@cam.ac.uk)

<sup>b</sup>Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, United Kingdom.

#### Materials

HPLC or PAGE-purified oligonucleotides were purchased from Sigma-Aldrich. Duplex DNA samples were prepared by reconstituting oligonucleotides in NH<sub>4</sub>OAc (200 mM, pH 7.0), heating complementary oligonucleotides at 95 °C for 10 min followed by slow cooling to room temperature, and then buffer-exchanging once more into NH<sub>4</sub>OAc (200 mM, pH 7.0). The sequences of the oligonucleotides used were (F and R denote forward and reverse complementary sequences, respectively):

DNA<sub>106</sub>F: 5'- CGG TCA TGG ATC CAC GCT ATC AAC GTA ATG TCG AGG CCG TCA ACG AGA TGT CGA CAC TAT CGA CAC GTA GTA AGC TGC CAG GGT GAC CAC CTC CGC GGC CAG TCA G -3'

DNA<sub>106</sub>R: 5'- CTG ACT GGC CGC GGA GGT GGT CAC CCT GGC AGC TTA CTA CGT GTC GAT AGT GTC GAC ATC TCG TTG ACG GCC TCG ACA TTA CGT TGA TAG CGT GGA TCC ATG ACC G -3'

DNA<sub>62</sub>F: 5' - GAT CCA CGC TAT CAA CGT AAT GTC GAG GCC GTC AAC GAG ATG TCG ACA CTA TCG ACA CGT AG -3'

DNA<sub>62</sub>R: 5' - CTA CGT GTC GAT AGT GTC GAC ATC TCG TTG ACG GCC TCG ACA TTA CGT TGA TAG CGT GGA TC -3'

DNA<sub>37</sub>F: 5'- ATC AAC GTA ATG TCG AGG CCG TCA ACG AGA TGT CGA C-3'

DNA<sub>37</sub>R: 5'- GTC GAC ATC TCG TTG ACG GCC TCG ACA TTA CGT TGA T -3'

DNA<sub>36</sub>F: 5' - GTA ATG TCG AGG CCG TCA ACG AGA TGT CGA CAC TAT -3'

DNA<sub>36</sub>R: 5' - ATA GTG TCG ACA TCT CGT TGA CGG CCT CGA CAT TAC -3'

DNA<sub>L311</sub>F: 5' - GAT CCA CGC TAT CAA CGT AAT GTC GAG GCC G -3'

DNA<sub>L31</sub>R: 5' - CGG CCT CGA CAT TAC GTT GAT AGC GTG GAT C -3'

DNA<sub>R31</sub>F: 5' - TCA ACG AGA TGT CGA CAC TAT CGA CAC GTA G -3'

DNA<sub>R31</sub>R: 5' - CTA CGT GTC GAT AGT GTC GAC ATC TCG TTG A -3'

### ***Protein expression and purification***

*Mycobacterium tuberculosis ethR* gene was cloned into a PHAT5 expression vector (*Anal. Biochem.* 1996, **236**, 371) having BamHI and EcoRI restriction sites. Expression and purification of His<sub>6</sub>-EthR protein was based on the protocol described in our previous work (*Biochem. J.* 2014, **458**, 387). *Escherichia* BL21 (DE3) strain (Novagen) containing the above plasmids were grown overnight at 37 °C in LB media for the primary culture. 25 mL of this culture was used to inoculate 1 L each of fresh LB media containing ampicillin (100 µg/mL) and was grown at 37 °C at 220 rpm until the culture reached an optical density ( $A_{600\text{nm}}$ ) of 0.6–0.8. The expression of the recombinant construct was induced by IPTG (0.5–1 mM) and the culture was further allowed to incubate at 37 °C for 3 h. The cells were harvested by centrifugation (4200 g, 20 min at 4 °C) and resuspended in lysis buffer [50 mM Hepes (pH 7.5), 150 mM NaCl] supplemented with EDTA-free protease inhibitor cocktail (Roche). The resuspended cells were lysed by sonication (Brason) using 10 pulses of 35% amplitude at intervals of 30 s each and the lysate was clarified by centrifugation (26,000 g for 1 h at 4 °C). The supernatant was passed-through a pre-equilibrated (with lysis Buffer) 5 mL HiTrap IMAC Fast Flow column (GE Healthcare), charged with Ni<sup>2+</sup>, at a flow rate of 3 mL/min. The column was washed with 50ml of wash buffer [50 mM Hepes (pH 7.5), 150 mM NaCl and 20 mM Imidazole] and the bound protein was eluted using elution buffer [50 mM Hepes (pH 7.5), 150 mM NaCl and 250 mM imidazole]. The protein thus obtained was further subjected to size-exclusion chromatography (Superdex 200 16/60, GE Healthcare) and fractions containing pure, homogenous EthR protein were pooled and concentrated (4500 g at 4 °C) using 10 kDa Amicon Ultra® centrifugal concentrators.

### ***Native nanoESI-MS***

Spectra were recorded on a Synapt HDMS mass spectrometer (Waters) modified for studying high masses. EthR and DNA samples were exchanged into NH<sub>4</sub>OAc (200 mM, pH 7.0) solution using Micro Bio-Spin 6 chromatography columns (Bio-Rad). For complexes, buffer-exchanged EthR and DNA were incubated together at the indicated concentrations for at least 15 min at room temperature before measurement. 2.5 µL of sample solution was injected into a borosilicate emitter (Thermo Scientific) for sampling. Instrument conditions were optimized to enhance ion desolvation while minimizing dissociation of macromolecular complexes. Typical conditions were capillary voltage 1.8–2.0 kV, sample cone voltage 100 V, extractor

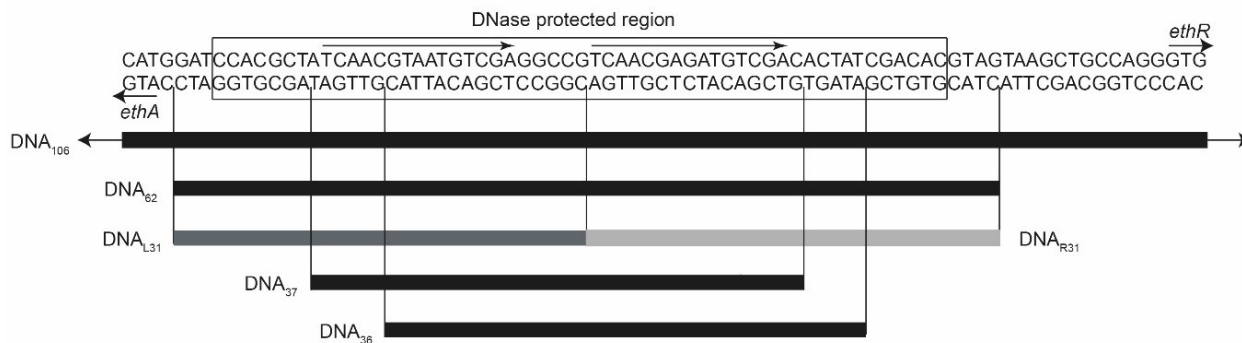
cone voltage 1 V, trap collision voltage 30 V, transfer collision voltage 30 V, source temperature 20 °C, backing pressure 5 mbar, trap pressure  $3\text{--}4 \times 10^{-2}$  mbar, IMS ( $\text{N}_2$ ) pressure  $5\text{--}6 \times 10^{-1}$  mbar and TOF pressure  $7\text{--}8 \times 10^{-7}$  mbar. Spectra were calibrated externally using cesium iodide. Data acquisition and processing were performed using MassLynx 4.1 (Waters).

Previous reports had described difficulties with achieving a stable electrospray with samples containing protein-DNA complexes (*Rapid Commun. Mass Spectrom.*, 2002, **16**, 1723; *J. Mol. Biol.*, 2013, **425**, 4790). Therefore, an extensive optimization campaign was carried out to determine the ideal parameters for detecting EthR-DNA complexes. A concentration of 200 mM of  $\text{NH}_4\text{OAc}$  gave the highest quality spectra in comparison to 20, 50 or 500 mM of  $\text{NH}_4\text{OAc}$  (Fig. S2). The addition of  $\text{Mg}^{2+}$  ions (500  $\mu\text{M}$ ) did not significantly increase the proportion of protein-DNA complexes (Fig. S3). Higher  $\text{Mg}^{2+}$  concentrations were not examined because non-volatile ions exert a detrimental effect on the quality of native mass spectra. The presence or absence of  $\text{Mg}^{2+}$  ions also did not influence the extent of EthR-DNA formation in an in-house SPR assay. Monitoring protein-DNA complex formation as a function of incubation time indicated that equilibrium was reached rapidly (<5 min) (Fig. S4). Additionally, spectra of protein-DNA complexes were not improved by buffer-exchanging a mixture of the EthR and the DNA together (Fig. S5). Finally, the quality of DNA-only mass spectra could be improved by buffer-exchanging duplexes into  $\text{NH}_4\text{OAc}$  after annealing, although a second buffer exchange step led to no further improvement in spectral quality (Fig. S6).

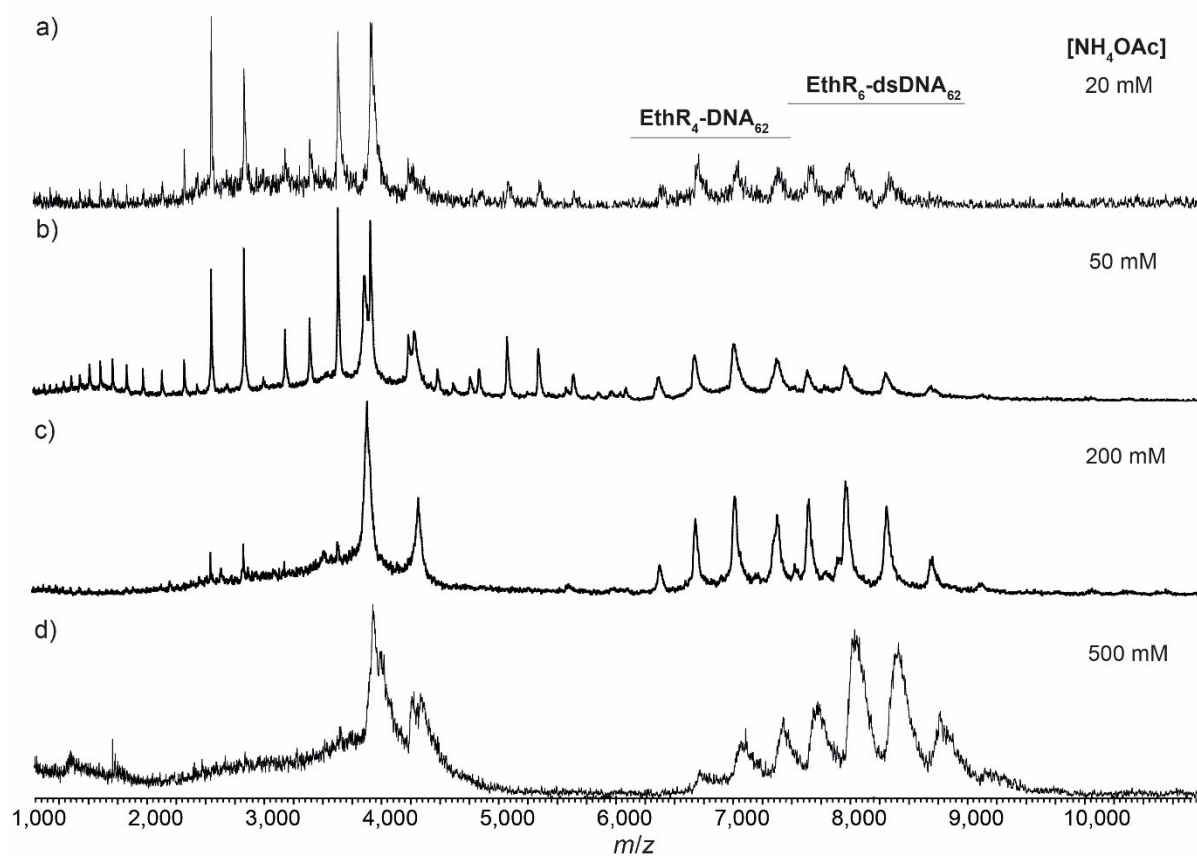
### ***Isothermal titration calorimetry (ITC)***

EthR and  $\text{DNA}_{62}$  were separately buffer-exchanged into a system of Tris-HCl (10 mM, pH 7.5), NaCl (200 mM),  $\text{MgCl}_2$  (2 mM) and EDTA (0.1 mM) using Vivaspin 500 centrifugal concentrators with a 10K MWCO (Sartorius). EthR was titrated into  $\text{DNA}_{62}$  in the reaction cell using a MicroCal Auto-iTC200 isothermal titration calorimeter (Malvern). The titration experiment was performed at 25 °C, and consisted of 20 injections of 2  $\mu\text{L}$  every 120 s, with stirring at 1,000 rpm. Data were fitted to a bespoke model ( $\text{EthR}_2 + \text{DNA}_{62} \rightleftharpoons \text{EthR}_2\text{-DNA}_{62} \rightleftharpoons \text{EthR}_4\text{-DNA}_{62} \rightleftharpoons \text{EthR}_6\text{-DNA}_{62}$ ) and analysed using the AFFINImeter software (S4SD). Global fitting of independently conducted ITC experiments was performed to improve the robustness of the fit. The concentrations of EthR and  $\text{DNA}_{62}$  in the two experiments are indicated in the table below.

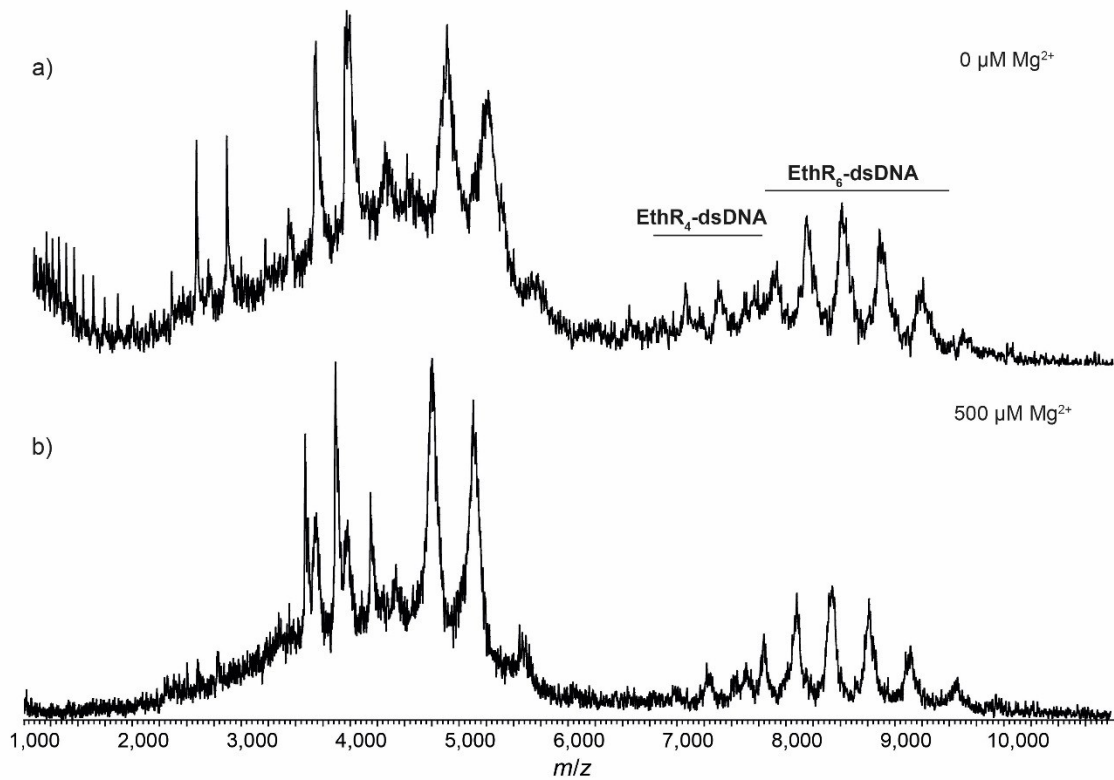
<b>Experiment</b>	<b>[EthR<sub>2</sub>] / μM</b>	<b>[DNA<sub>62</sub>] / μM</b>
1	418	10.7
2	548	15.7



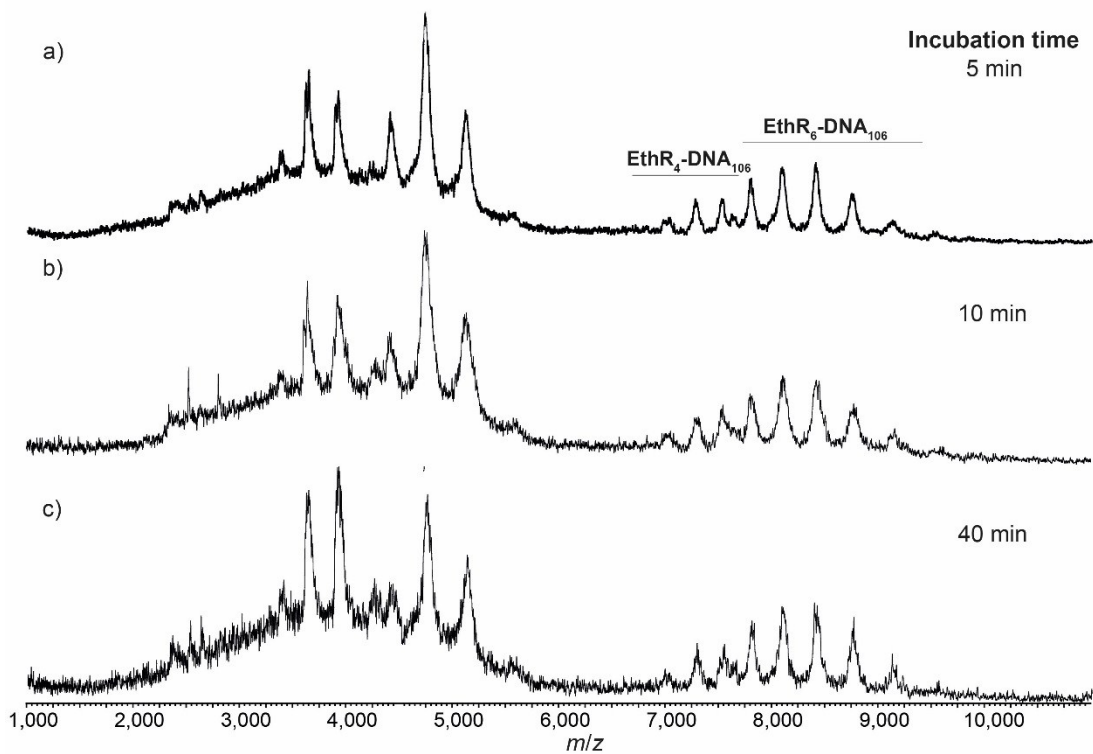
**Figure S1.** Sequence of the *ethA-R* intergenic region. The positions of the DNA sequences used in this work are shown. The two arrows in the DNase protected region indicate direct repeats.



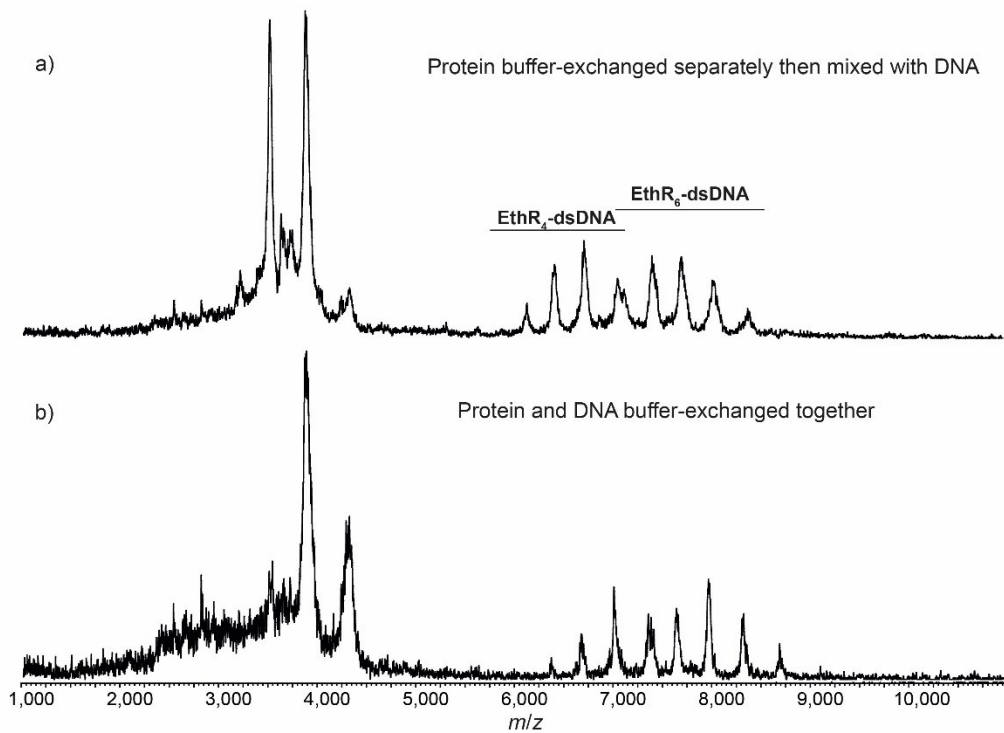
**Figure S2.** Native mass spectra of EthR (20  $\mu$ M) and DNA<sub>62</sub> (2.5  $\mu$ M) in NH<sub>4</sub>OAc (pH 7.0) of concentration a) 20 mM, b) 50 mM, c) 200 mM, d) 500 mM.



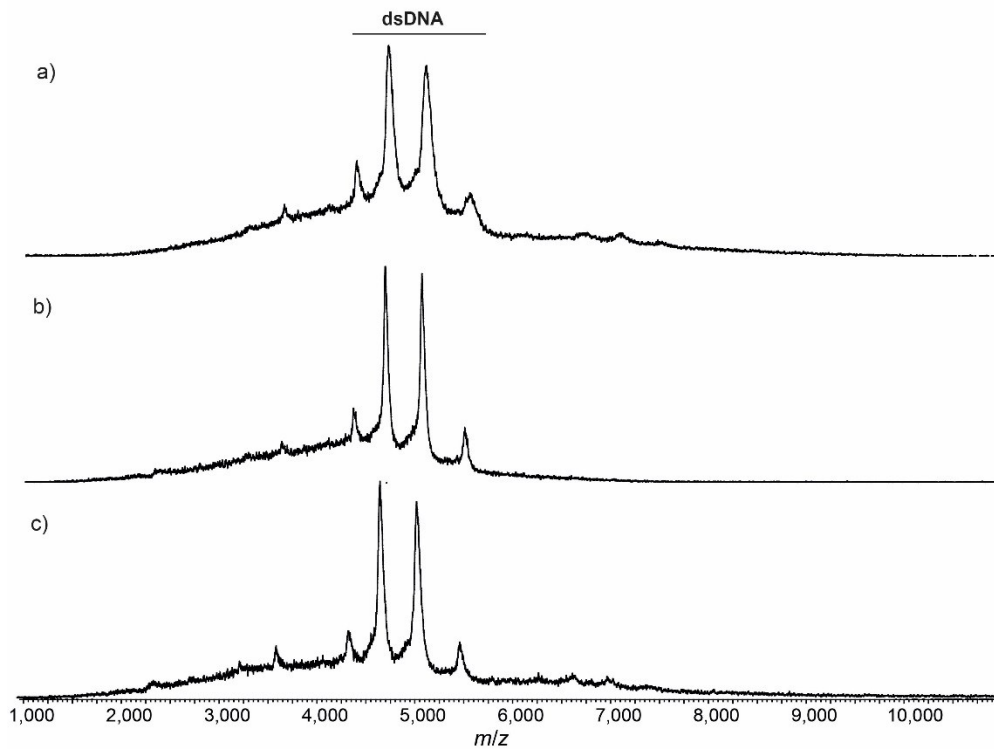
**Figure S3.** Native mass spectra of EthR (20  $\mu M$ ) and DNA<sub>106</sub> (2.5  $\mu M$ ) in NH<sub>4</sub>OAc (200 mM, pH 7.0) in the a) absence and b) presence of  $Mg^{2+}$  ions (500  $\mu M$ ).



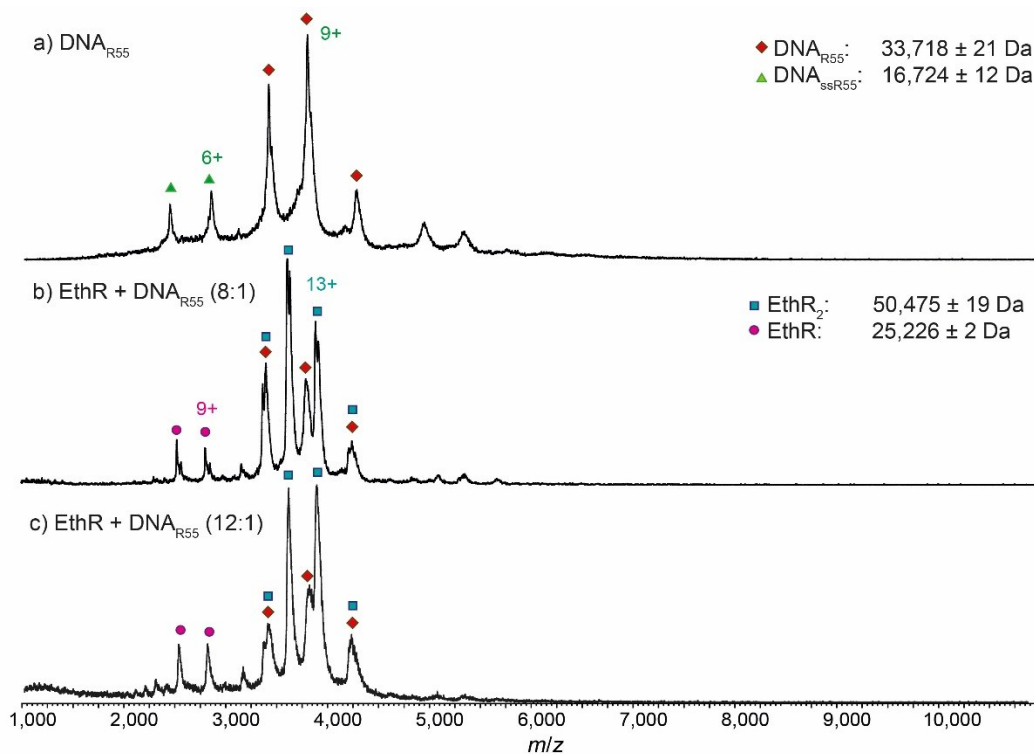
**Figure S4.** Native mass spectra of EthR (20  $\mu M$ ) and DNA<sub>106</sub> (2.5  $\mu M$ ) in NH<sub>4</sub>OAc (200 mM, pH 7.0) after incubation for a) 5 min, b) 10 min, and c) 40 min.



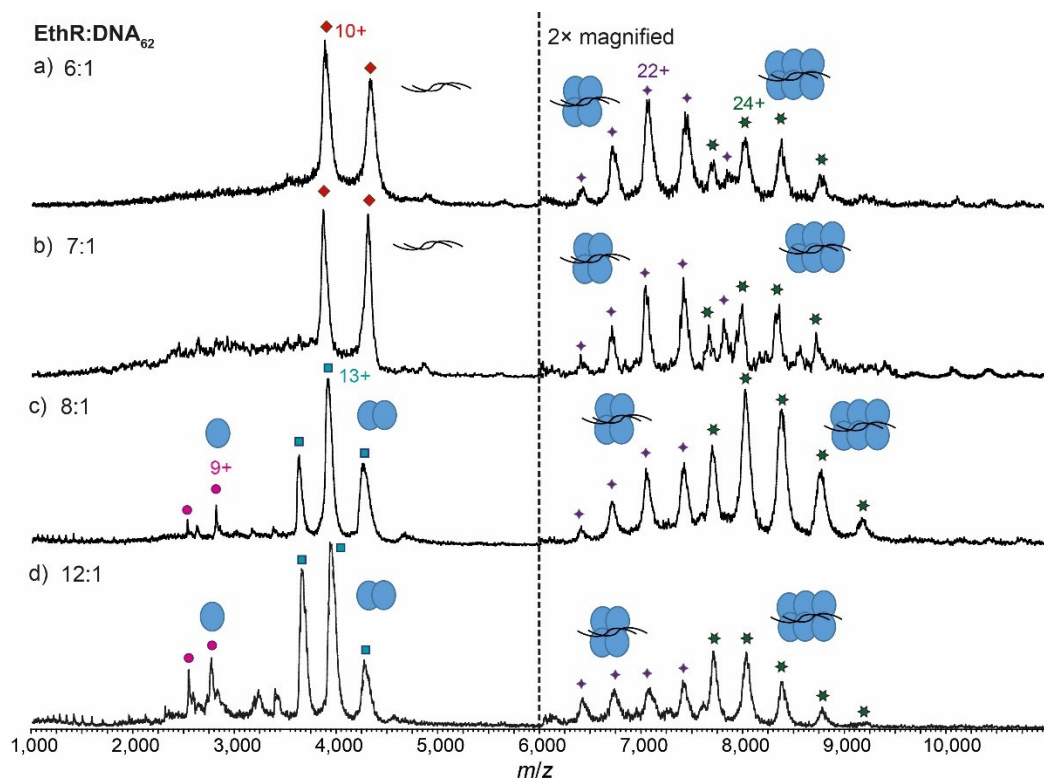
**Figure S5.** Native mass spectra of EthR (20  $\mu\text{M}$ ) and DNA<sub>62</sub> (2.5  $\mu\text{M}$ ) in NH<sub>4</sub>OAc (200 mM, pH 7.0). In a), buffer-exchanged protein was incubated with DNA before measurement. In b), protein was incubated with DNA, and the mixture was buffer-exchanged before measurement.



**Figure S6.** a) Native mass spectra of DNA<sub>106</sub> (10  $\mu\text{M}$ ) in NH<sub>4</sub>OAc (200 mM, pH 7.0), and after buffer-exchanging into NH<sub>4</sub>OAc (200 mM, pH 7.0) b) once or c) twice.

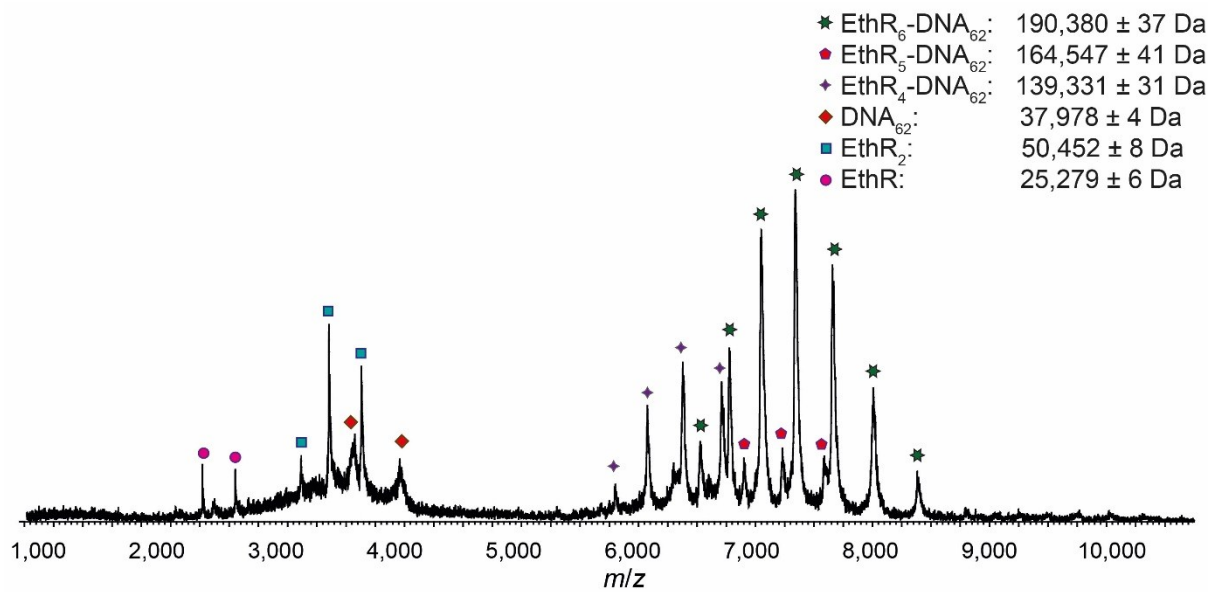


**Figure S7.** Native mass spectra a) DNA<sub>R55</sub> (2.5 μM) and b) a mixture of EthR (20 μM) and DNA<sub>R55</sub> (2.5 μM) in an 8:1 ratio, and c) a mixture of EthR (30 μM) and DNA<sub>R55</sub> (2.5 μM) in a 12:1 ratio. All spectra were recorded in NH<sub>4</sub>OAc (200 mM, pH 7.0).

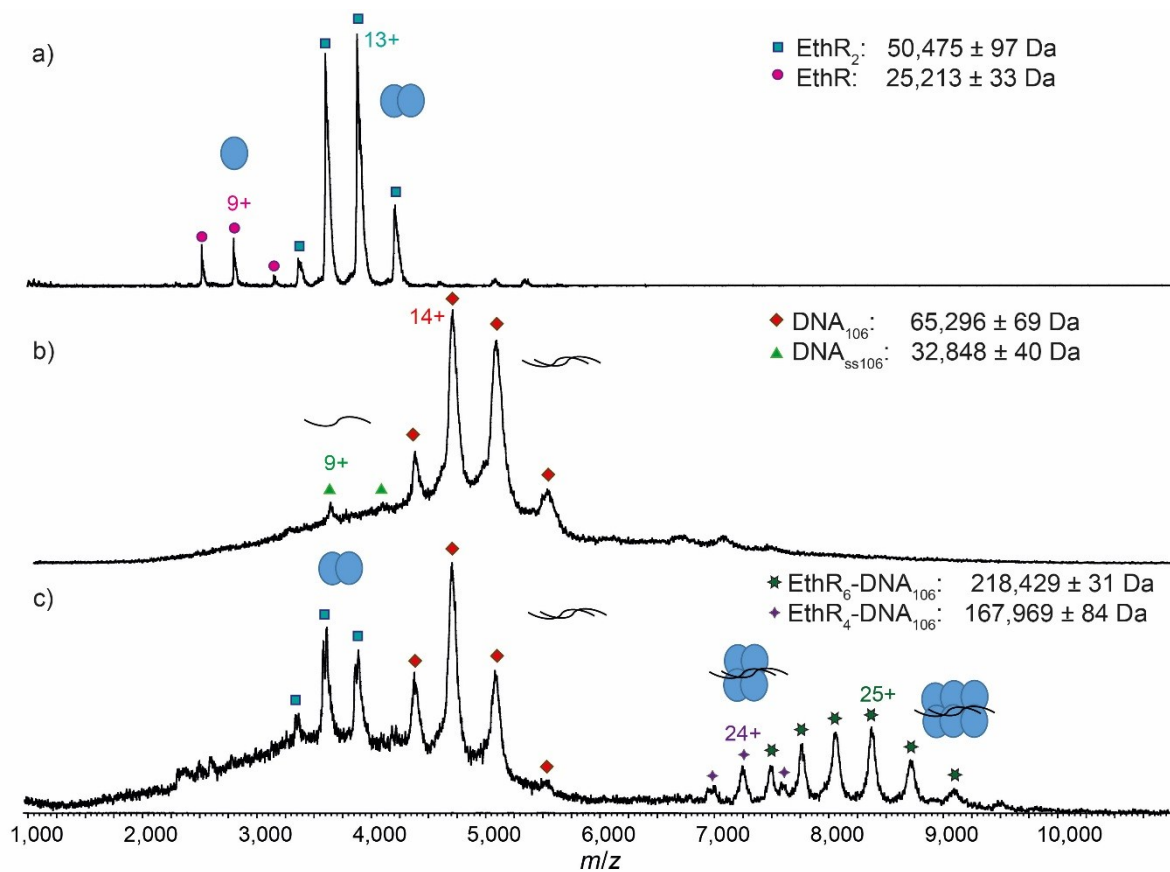


**Figure S8.** Native mass spectra of a mixture of EthR and DNA<sub>62</sub> in a) 6:1 ratio, b) 7:1 ratio, c) 8:1 ratio or c) 12:1 ratio. The concentration of DNA<sub>62</sub> was held at 2.5 μM. All spectra were recorded in NH<sub>4</sub>OAc (200 mM, pH 7.0). The protein-DNA peaks have been magnified by 2× for clarity.

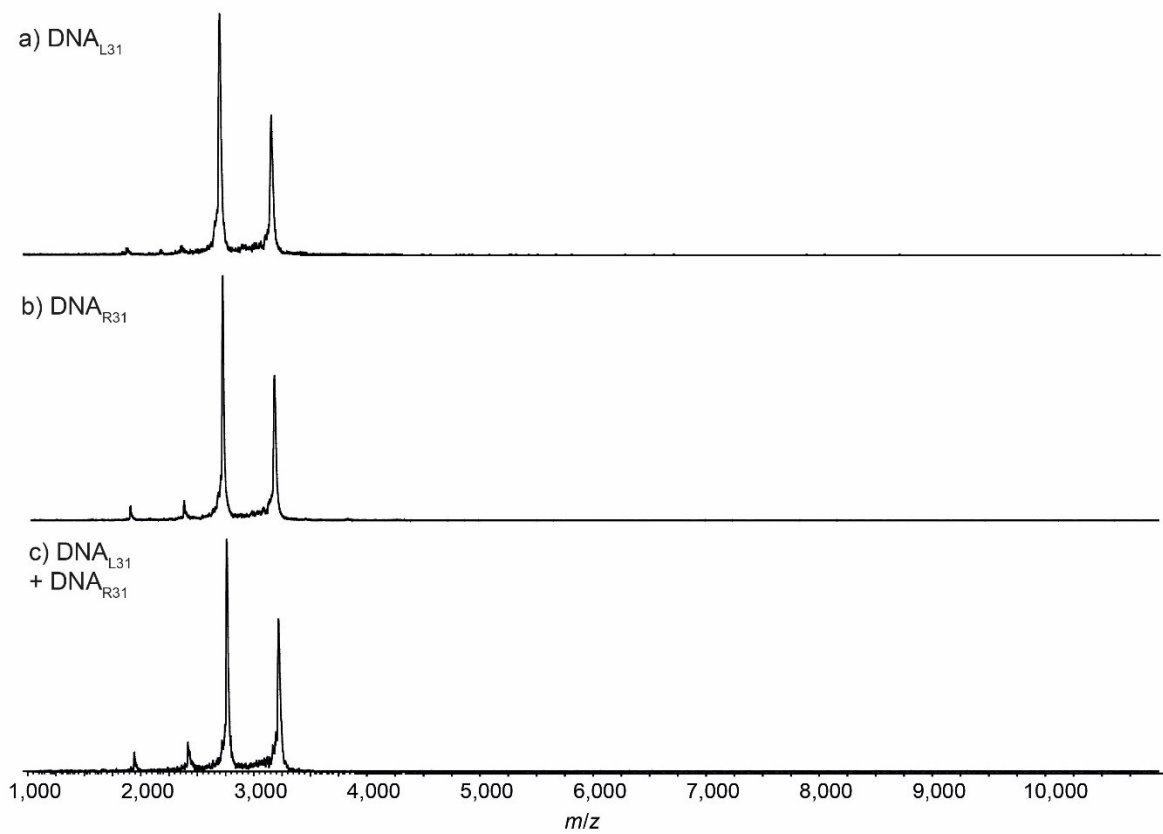




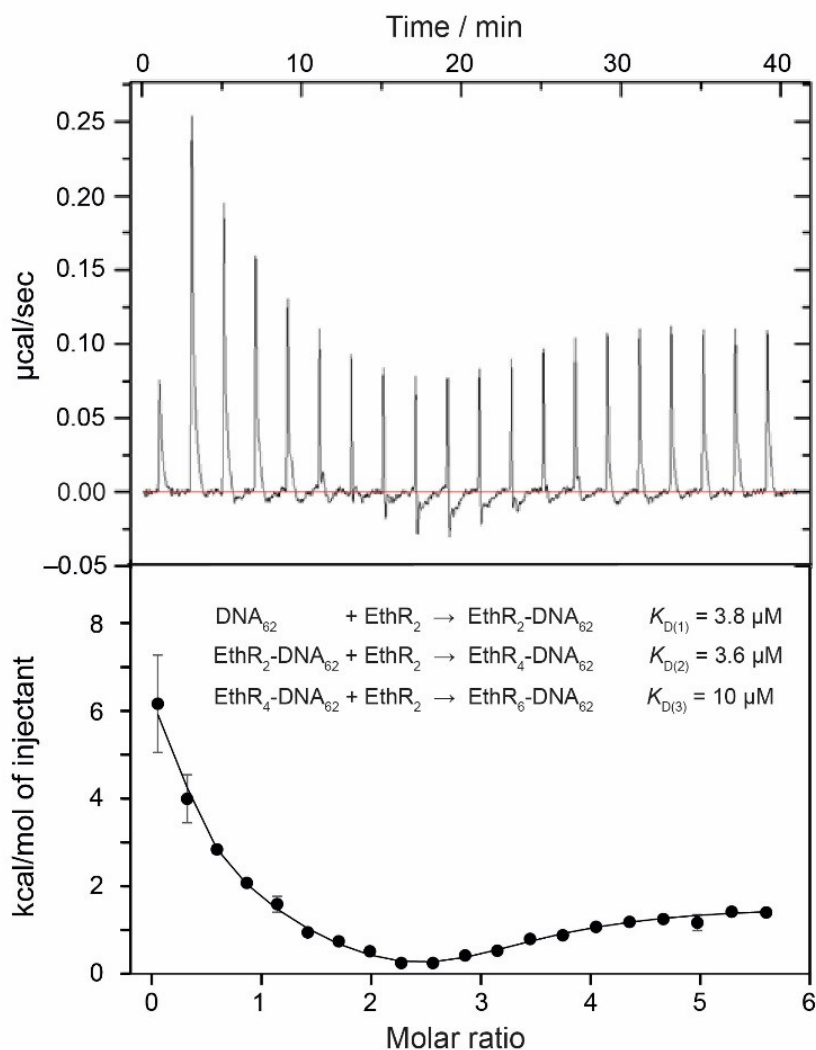
**Figure S9.** Native mass spectra of a mixture of EthR (20  $\mu$ M) and DNA<sub>62</sub> (2.5  $\mu$ M) in NH<sub>4</sub>OAc (200 mM, pH 7.0).



**Figure S10.** Native mass spectra a) EthR (20  $\mu$ M), b) DNA<sub>106</sub> (2.5  $\mu$ M) and c) a mixture of EthR (20  $\mu$ M) and DNA<sub>106</sub> (2.5  $\mu$ M). All spectra were recorded in NH<sub>4</sub>OAc (200 mM, pH 7.0).



**Figure S11.** Native mass spectra of a) DNA<sub>L31</sub> (2.5  $\mu$ M), b) DNA<sub>R31</sub> (2.5  $\mu$ M) and c) a mixture of DNA<sub>L31</sub> (2.5  $\mu$ M) and DNA<sub>R31</sub> (2.5  $\mu$ M). All spectra were recorded in NH<sub>4</sub>OAc (200 mM, pH 7.0).



**Figure S12.** ITC data for the titration of EthR into DNA<sub>62</sub> showing the thermogram (upper panel) and binding isotherm (lower panel). Integrated data were fit with a stoichiometric equilibrium binding model. Moles of EthR are given as the dimer. ITC data were fit to an equilibrium binding model that assumes the stepwise formation of the EthR<sub>6</sub>-DNA complex from EthR dimers, as indicated in the lower panel along with the associated equilibrium dissociation constants for the binding of the first ( $K_{D(1)}$ ), second ( $K_{D(2)}$ ) and third ( $K_{D(3)}$ ) EthR dimer to the DNA.