Supplementary Information

Single-molecule DNA unzipping reveals asymmetric modulation of a transcription factor by its binding site sequence and context

Sergei Rudnizky[1,†], Hadeel Khamis[1,2,†], Omri Malik[1,3], Allison H. Squires[4,#], Amit Meller[3,4,5], Philippa Melamed[1,3] and Ariel Kaplan[1,3,*]

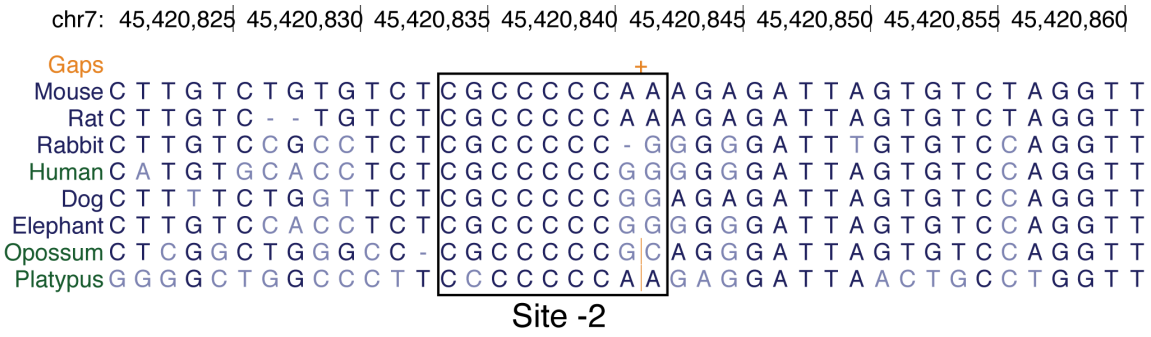
[1]Faculty of Biology, Technion – Israel Institute of Technology, Haifa (32000), Israel, [2]Faculty of Physics, Technion – Israel Institute of Technology, Haifa (32000), Israel, [3]Russell Berrie Nanotechnology Institute, Technion – Israel Institute of Technology, Haifa (32000), Israel, [4]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA and [5]Faculty of Biomedical Engineering, Technion – Israel Institute of Technology, Haifa (32000), Israel

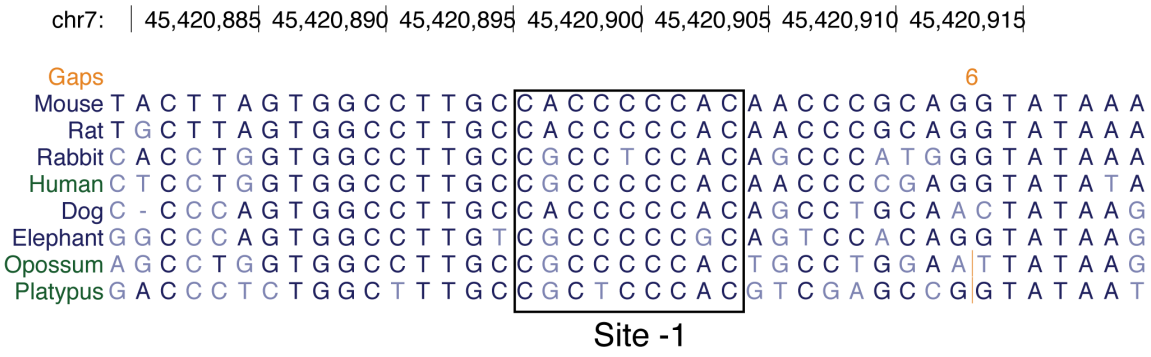
* To whom correspondence should be addressed. Tel: +972-77-8871907; Fax: +972-77-8871908; Email: akaplanz@technion.ac.il

† These authors contributed equally to the paper as first authors

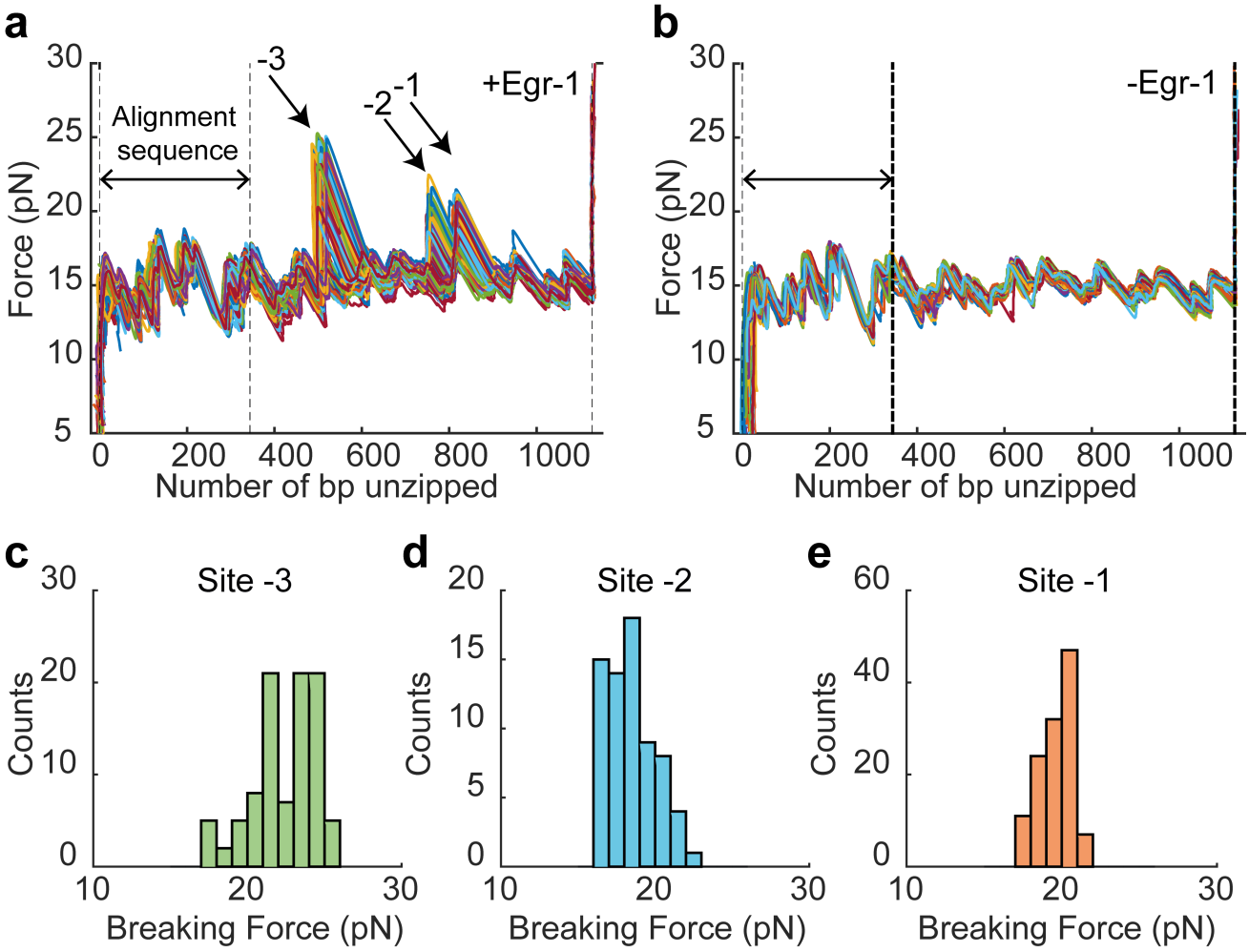
#Present address: Department of Chemistry, Stanford University, Stanford, CA 94305

SUPPLEMENTARY FIGURES





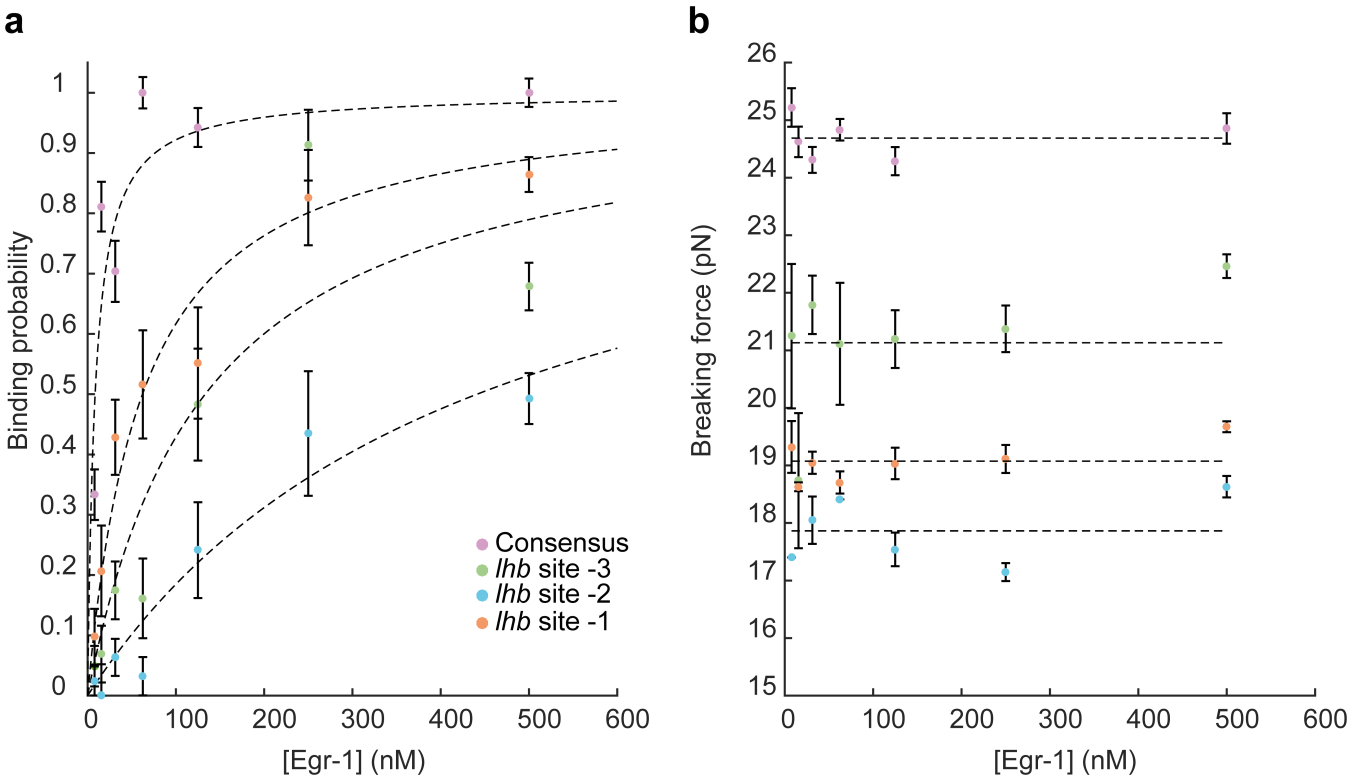
Supplementary Figure S1: Binding sites for Egr-1 at the promoter of *Lhb* are evolutionary conserved. Alignment of the proximal *Lhb* promoters from different species. The conserved Egr-1 response elements are boxed. Nucleotides that differ between the species are shaded. The data was generated with the Genome Browser software, http://genome.ucsc.edu/.



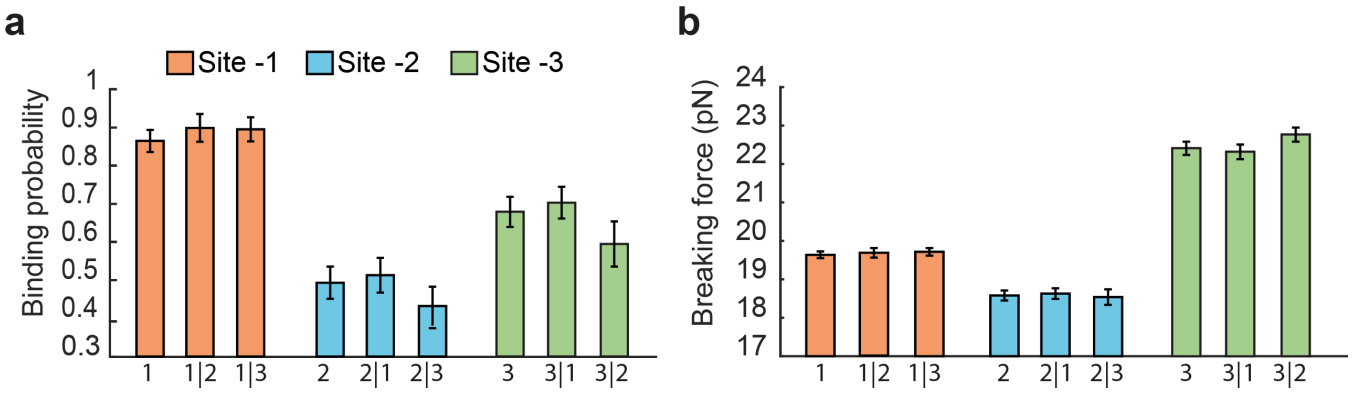
Supplementary Figure S2: Characterization of binding probability and breaking force for sites -1,-2 and -3 in *Lhb*. (a-b) All the collected traces for forward unzipping of *Lhb* in the presence (a) and absence (b) of Egr-1 (n=140 and 48, respectively). (c-e) Histograms for the breaking force for sites -3 (c), -2 (d) and -1 (e).

lhb site -3	....5'-G C A A A T T T G G A [ggcccactc] G T C A G A A C C T A-3'...
lhb site -2	....5'-T G T C T G T G T C T [cgcccccaa] A G A G A T T A G T G-3'...
lhb site -1	....5'-A G T G G C C T T G C [caccccccac] A A C C C G C A G G T-3'...
C1	....5'-T A T C T G A C A C G [Binding site] A C T A G G G A G T A-3'...
C2	....5'-T C A C T G C C G G C [Binding site] G A T G T A T A T A T-3'...

Supplementary Figure S3: DNA context surrounding the 9 bp core binding sites probed in the unzipping experiments. 11 base pairs surrounding the naturally occurring sites on the *Lhb* promoter (*Lhb* sites -1, -2 and -3) or supplied by the C1 or C2 contexts where the core binding sites are integrated.

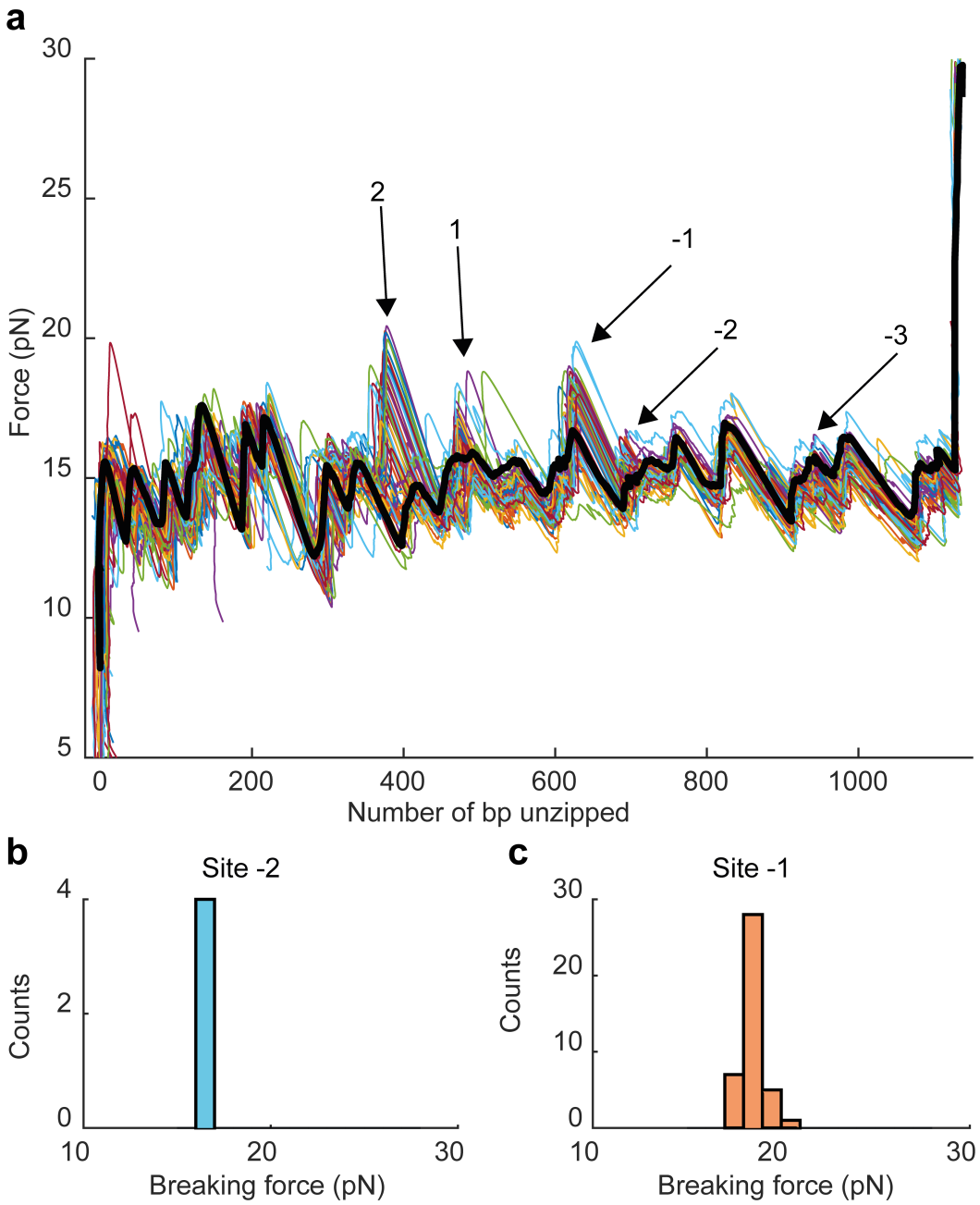


Supplementary Figure S4: The binding probability depends on Egr-1 concentration, but the breaking force is concentration independent. Multiple unzipping experiments were performed for the *Lhb* binding sites and for the consensus sequence integrated into the C1 DNA context, as a function of the concentration of Egr-1. (a) The binding probability and (b) mean breaking force data presented for sites -3/-2/-1 and consensus. The data is analyzed as in Fig 1. The dashed lines are hyperbolic (a) and linear (b) fits. The number of events for each case is shown in Supplementary Table 4.

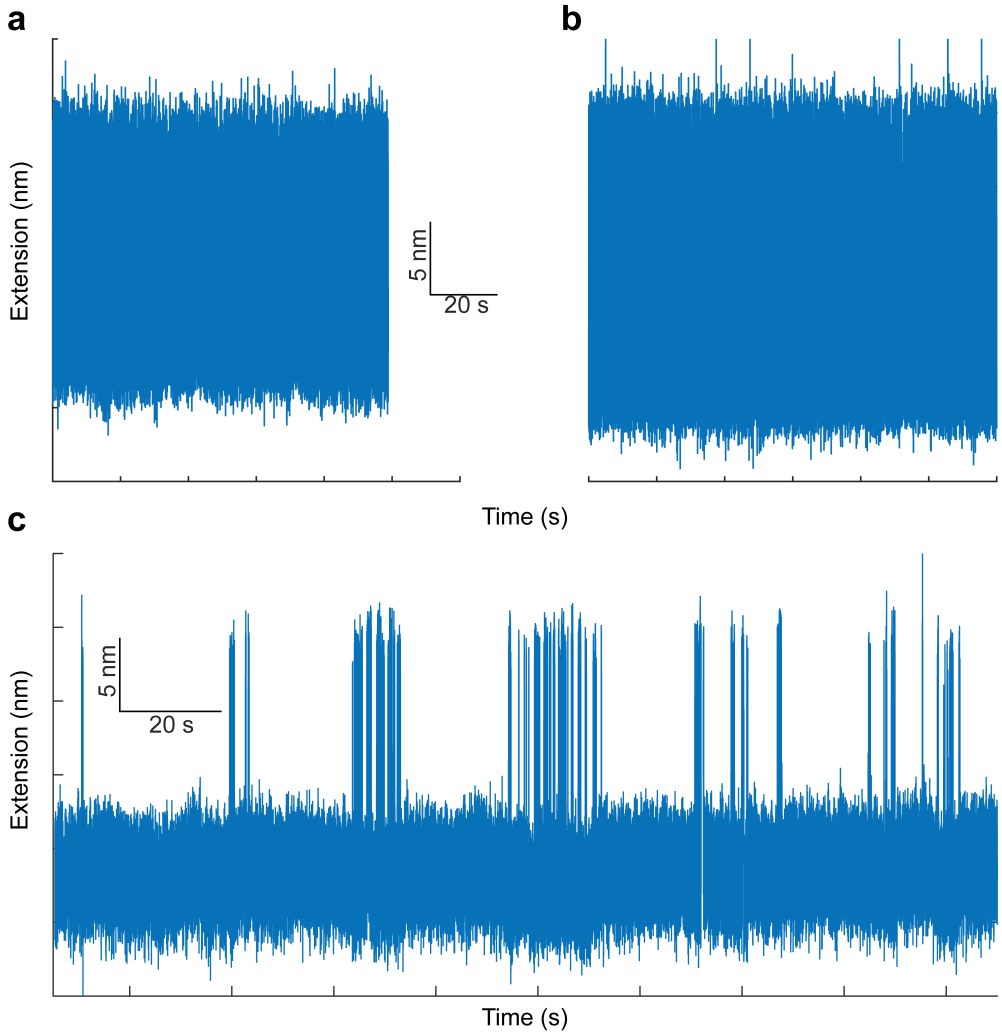


Supplementary Figure S5: Binding of Egr-1 to a given site is independent of the occupancy of other sites *in cis*. Binding probabilities (a) and mean breaking force (b) are shown for sites -1, -2 and -3 on *Lhb* DNA, both for the complete data set as well as for the conditional subsets (*r*|*s* indicates binding to site *r*, provided that site *s* was bound too). The data is analyzed and presented as in Fig 1. *n*=140.

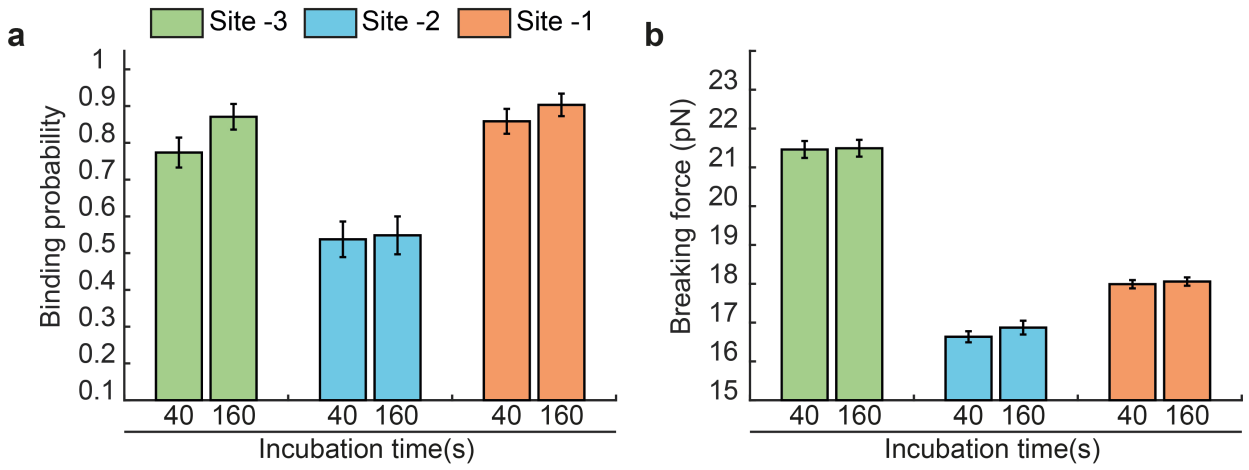
Supplementary Figure S6: Constructs used to characterize the extent of the flanking sequence effect. (a-b) Constructs used to perturb the flanking sequence at the ZF1 (a) or ZF3 (b) side of the core binding site. Mutagenized bases are shown in blue.



Supplementary Figure S7: Characterization of binding probability and breaking force by reverse-unzipping *Lhb*. (a) All the collected traces for reverse unzipping of *Lhb* in the presence of Egr-1 (n = 63). The position of sites -1, -2 and -3 is indicated. Sites 1 and 2, located in the gene body, were only observed in reverse unzipping. A representative trace in the absence of Egr-1 is shown in black. (b) Histograms for the breaking force for sites -2 and -1.



Supplementary Figure S8: Dissociation time measurements. (a-b) No suppression of fluctuations is observed in the absence of Egr-1, both for forward (a) and reverse (b) unzipping. (c) Multiple events of binding and dissociation are observed if the construct is kept at the Egr-1 flow channel. Data shown for reverse unzipping.



Supplementary Figure S9: The incubation time is long enough to allow equilibration. (a-b) Unzipping force-extension curves of *Lhb* in the presence of 500 nM Egr-1, for two different incubation times (defined as the time between successive unzipping cycles of a specific binding site). (a) Binding probability, calculated as the number of binding events out of the total number of DNA unzipping cycles, for each of the three Egr-1 binding sites. Data shown as fraction ± s.e., *n*=106,93. (b) Breaking force for three Egr-1 binding sites located on the *Lhb* promoter. Data shown as mean ± s.e., n=106,93. No significant differences were detected.

SUPPLEMENTARY TABLES

Supplementary Table S1: Number of experiments for the data in Figs. 2 and 5

Context	Site -1	Site -2	Site -3	Consensus
C2	41	63	33	64
C1	213	47	78	41
C1 (reverse)	-	-	-	57
Lhb (forward)	140	140	140	-
Lhb (reverse)	63	63	63	-

Supplementary Table S2: Number of experiments for the data in Fig. 3

Mutation	Consensus	Site -1
3'-C	162	111
3'-G	183	331
3'-T	178	482
5'-A	83	180
5'-C	76	103
5'-T	51	120

Supplementary Table S3: Number of experiments for the data in Fig. 4

Nearest perturbed bp	Site -1	Nearest perturbed bp	Site -1
10	58	-1	73
11	26	-2	140
14	106	-3	75
15	108	-4	109
16	57	-5	86
18	218	-6	130
19	70	-8	109
20	72		

Supplementary Table S4: Number of experiments for the data in Supplementary Fig. 4

Concentration (nM)	8	17	31	62	125	250	500
lhb	41	29	63	31	29	23	140
Consensus	126	90	81	37	52	-	41

Supplementary Table S5: Calculated p-values for the data in Fig. 2

Breaking force

	Site -1			Site -2			Site -3			Consensus	
	C2	C1	*Lhb*	C2	C1	*Lhb*	C2	C1	*Lhb*	C2	C1
C2		0.02	0.1		0.7	2E-08		0.7	1E-07		0.0007
C1			0.001			9E-07			8E-08		
Lhb											

Binding probability

	Site 1			Site 2			Site 3			Consensus	
	C2	C1	*Lhb*	C2	C1	*Lhb*	C2	C1	*Lhb*	C2	C1
C2		0.01	0.007		5E-02	9E-02		0.02	0.8		0.1
C1			0.6			4E-04			0.001		
Lhb											

Supplementary Table S6: Calculated p-values for the data in Fig. 3

Binding probability, consensus

Mutation	C1	3'-C	3'-G	3'-T	5'-A	5'-C	5'-T
C1		0.2	0.5	0.007	0.3	0.01	0.1
3'-C			0.3	0.04	0.8	0.07	0.4
3'-G				0.004	0.5	0.02	0.2
3'-T					0.05	0.6	0.5
5'-A						0.06	0.4
5'-C							0.3
5'-T							

Binding probability, site -1

Mutation	C1	3'-C	3'-G	3'-T	5'-A	5'-C	5'-T
C1		1E-09	1E-09	1E-09	3E-05	0.005	0.5
3'-C			1E-01	0.1	1E-22	2E-14	2E-07
3'-G				0.7	1E-23	3E-14	6E-07
3'-T					1E-24	2E-14	6E-07
5'-A						0.4	5E-06
5'-C							0.001
5'-T							

Breaking force, consensus

Mutation	C1	3'-C	3'-G	3'-T	5'-A	5'-C	5'-T
C1		0.03	0.02	0.01	0.3	0.7	0.2
3'-C			0.4	0.7	0.05	0.001	0.1
3'-G				0.6	0.02	0.0005	0.08
3'-T					1E-02	0.0001	7E-02
5'-A						0.2	0.7
5'-C							0.1
5'-T							

Breaking force, site -1

Mutation	C1	3'-C	3'-G	3'-T	5'-A	5'-C	5'-T
C1		4E-11	7E-08	2E-43	1E-07	6E-08	4E-10
3'-C			0.04	1E-04	0.01	0.09	2E-01
3'-G				2E-18	0.6	0.7	2E-01
3'-T					8E-21	7E-13	2E-10
5'-A						0.4	0.1
5'-C							0.4
5'-T							

Supplementary Table S7: Calculated p-values for the data in Fig. 4

Breaking force, site -1, ZF1

Nearest perturbed bp	C1	-1	-2	-3	-4	-5	-6	-8
C1		5E-10	2E-08	0.0002	0.04	0.08	0.9	0.1
-1			0.1	5E-02	0.0003	0.0001	7E-08	1E-05
-2				0.3	0.005	0.002	1E-06	0.0003
-3					0.1	0.07	0.001	0.03
-4						0.7	0.07	0.5
-5							0.1	0.7
-6								0.1
-8								

Binding probability, site -1, ZF1

Nearest perturbed bp	C1	-1	-2	-3	-4	-5	-6	-8
C1		0.003	0.08	0.01	0.9	0.4	0.6	0.006
-1			5E-05	0.5	0.005	0.03	0.001	0.4
-2				0.0003	0.1	0.02	0.2	3E-05
-3					0.02	0.1	0.006	0.9
-4						0.4	0.5	0.01
-5							0.2	0.08
-6								0.002
-8								

Breaking force, site -1, ZF3

Nearest perturbed bp	C1	10	11	14	15	16	18	19	20
C1		3E-41	1E-26	1E-26	5E-28	2E-22	2E-34	2E-22	3E-29
10			0.001	5E-07	7E-07	1E-05	1E-20	2E-18	3E-22
11				0.1	0.1	0.2	9E-07	2E-07	2E-08
14					0.7	0.6	0.002	0.0001	0.0001
15						0.8	0.0004	3E-05	1E-05
16							0.001	0.0001	5E-05
18								0.09	0.1
19									0.6
20									

Binding probability, site -1, ZF3

Nearest perturbed bp	C1	10	11	14	15	16	18	19	20
C1		0.0001	0.9	0.6	0.5	0.4	0.4	0.9	0.4
10			3E-10	2E-09	2E-08	3E-07	7E-11	4E-10	4E-15
11				0.5	0.3	0.2	0.1	0.9	0.3
14					0.1	0.1	0.08	0.5	0.8
15						0.7	0.7	0.3	0.04
16							0.9	0.2	0.02
18								0.2	0.01
19									0.3
20									

Supplementary Table S8: Calculated p-values for the data in Fig. 5

	Site -1	Site -2	Site -3	Consensus
Breaking force	4E-11	0.018	-	4E-22
Binding probability	4E-04	3E-09	3E-19	0.48

Supplementary Table S9: Oligonucleotides used to synthesize the *Lhb* constructs.

Name	Sequence
mLh -517F BglI	ATGGCCTTGCCGGCGACACACCCCTTACTTCCAGAG
mLh +246R DraIII	ATCACTGCGTGTCTACCCCTGACCTGGTTTTTC

Supplementary Table S10: Oligonucleotides used to synthesize the C1 and C2 constructs.

Common reverse primer	601 R DraIII HP side	ATCACTGCGTGGTCACCGATGGACCCTATACG
Forward primers for WT	cons_5_prime_601_WT	ATGGCCTTGCCGGCcgcccacgcGATGTATATATCTGACACGTGCCTGGAG ACTA
	site3_5_prime_601_WT	ATGGCCTTGCCGGCggcccactcGATGTATATATCTGACACGTGCCTGGAG ACTA
	site2_5_prime_601_WT	ATGGCCTTGCCGGCcgcccccaaGATGTATATATCTGACACGTGCCTGGAG ACTA
	site1_5_prime_601_WT	ATGGCCTTGCCGGCcacccccacGATGTATATATCTGACACGTGCCTGGAG ACTA
	cons_601_WT	ATGGCCTTGCCGGCGATGTATATATCTGACACGcgcccacgcACTAGGGAG TAATCCCCTTG
	site3_601_WT	ATGGCCTTGCCGGCGATGTATATATCTGACACGggcccactcACTAGGGAG TAATCCCCTTG
	site2_601_WT	ATGGCCTTGCCGGCGATGTATATATCTGACACGcgcccccaaACTAGGGAG TAATCCCCTTG
	site1_601_WT	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTAGGGAG TAATCCCCTTG
Forward primers for single mutations	site1_601_3flankT	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacTCTAGGGAG TAATCCCCTTG
	site1_601_3flankC	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacCTAGGGAGT AATCCCCTTG
	site1_601_3flankG	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacGCTAGGGAG TAATCCCCTTG
	site1_601_5flankT	ATGGCCTTGCCGGCGATGTATATATCTGACACTcacccccacACTAGGGAG TAATCCCCTTG
	site1_601_5flankC	ATGGCCTTGCCGGCGATGTATATATCTGACACCcacccccacACTAGGGAG TAATCCCCTTG
	site1_601_5flankA	ATGGCCTTGCCGGCGATGTATATATCTGACACAcacccccacACTAGGGAG TAATCCCCTTG
	cons_601_3flankT	ATGGCCTTGCCGGCGATGTATATATCTGACACGcgcccacgcTCTAGGGAG TAATCCCCTTG
	cons_601_3flankC	ATGGCCTTGCCGGCGATGTATATATCTGACACGcgcccacgcCCTAGGGAG TAATCCCCTTG
	cons_601_3flankG	ATGGCCTTGCCGGCGATGTATATATCTGACACGcgcccacgcGCTAGGGAG TAATCCCCTTG
	cons_601_5flankT	ATGGCCTTGCCGGCGATGTATATATCTGACACTcgcccacgcACTAGGGAG

		TAATCCCCTTG
	cons_601_5flankC	ATGGCCTTGCCGGCGATGTATATATCTGACACCcgcccacgcACTAGGGAG TAATCCCCTTG
	cons_601_5flankA	ATGGCCTTGCCGGCGATGTATATATCTGACACAcgcccacgcACTAGGGAG TAATCCCCTTG
Forward primers for multiple mutations	site1_5flanKO8	ATGGCCTTGCCGGCGATGTATATATATGACACGcacccccacACTAGGGAG TAATCCCCTTG
	site1_5flanKO6	ATGGCCTTGCCGGCGATGTATATATATAACACGcacccccacACTAGGGAG TAATCCCCTTG
	site1_5flanKO5	ATGGCCTTGCCGGCGATGTATATATATATCACGcacccccacACTAGGGAG TAATCCCCTTG
	site1_5flanKO4	ATGGCCTTGCCGGCGATGTATATATATATAACGcacccccacACTAGGGAG TAATCCCCTTG
	site1_5flanKO3	ATGGCCTTGCCGGCGATGTATATATATATATCGcacccccacACTAGGGAG TAATCCCCTTG
	site1_5flanKO2	ATGGCCTTGCCGGCGATGTATATATATATATAGcacccccacACTAGGGAG TAATCCCCTTG
	site1_5flanKO1	ATGGCCTTGCCGGCGATGTATATATATATATATcacccccacACTAGGGAG TAATCCCCTTG
	site1_3flanKO11	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTAGGGAG TTATCCCCTTGGCGGTTAAAA
	site1_3flanKO10	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTAGGGAG ATATCCCCTTGGCGGTTAAAA
	site1_3flanKO9	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTAGGGAT ATATCCCCTTGGCGGTTAAAA
	site1_3flanKO7	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTAGGTAT ATATCCCCTTGGCGGTTAAAA
	site1_3flanKO6	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTAGATAT ATATCCCCTTGGCGGTTAAAA
	site1_3flanKO5	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacACTATATAT ATATCCCCTTGGCGGTTAAAA
	site1_3flanKO2	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacAATATATAT ATATCCCCTTGGCGGTTAAAA
	site1_3flanKO1	ATGGCCTTGCCGGCGATGTATATATCTGACACGcacccccacTATATATAT ATATCCCCTTGGCGGTTAAAA

Supplementary Table S11: Oligonucleotides used to synthesize the dsDNA handles and alignement segment.

2k BioTg F	GATCTCCAGCCAGGAACTATTGA
2k bio R Nt.BvCI	GTGTCAGCTTGCCCCTCAGCGATGACCTCAGCATTTTTCGACCTGCTCTTCAGCA
2K DoubleDig2 bgl1F	ATGGCCTAGACGGCGAGCCTGGGTTTATAAGGGGAGCGGTGA
2k Dig Nb.BbVCI R	TCAGCTTGCCCCTCAGCGATGACCTCAGCAAGGACCAGCGTTTTGTTGAAA
Oligo 5' Dig AGA	AGGTGCCGGGACCACCCTGTAGA
Oligo 3' Dig/ 5' phosphate	ACAGGGTGGTCCCGGCACCT
Alignment F DraIII CAC	ATCACCACGTGGAATTCGATATCCCCGAGA
Alignment R DraIII	ATGCACGCAGTGACCATGGTGGTGTTTCCC
Hairpin oligo 5' phos	GACTTGAGGCAATTGCCTCAAGTCtgc