Tables with information used for DNA sequence preference study (protocol 2) for bzip, p22 and TATA box binding proteins.

Table 1. Information used for bZIP-DNA binding. We asked MELD-DNA to satisfy at least 5 groups in each collection. Groups where defined based on all the combinatorics of residues in DNA pairing with those in the protein. Each group had 3 possible restraints between the C_b in the protein and 3 atoms in DNA.

Collection 1 (Domain 1)					
	DNA-resid	**DNA atom**	**Prot-resid**	**Protein atom**	**Distance (Å)**
Groups (70)	121 to 125	P, OP1, OP2	4 to 17	C_b	8.0
Collection 2 (Domain 2)					
	DNA-resid	**DNA atom**	**Prot-resid**	**Protein atom**	**Distance (Å)**
Groups (75)	116 to 120	P, OP1, OP2	60 to 74	C_b	8.0

Table 2. Information used for the P22-DNA binding. Each group was evaluated best on one restraint and each collection was enforcing three groups according to the MELD-DNA strategy. Concretely, DNA Residue ID (DNA-resid), DNA atom, Protein Residue ID (Prot-Resid), Protein atom and distances (in Å) are indicated.

Collection 1 (Domain 1)					
	DNA-resid	**DNA atom**	**Prot-resid**	**Protein atom**	**Distance (Å)**
Group 1	21	O5'	49	C_b	8.0
	21	O5'	59	C_b	8.0
	21	O5'	70	C_b	8.0
	21	O5'	74	C_b	8.0
	21	O5'	75	C_b	8.0
Group 2	22	P	49	C_b	8.0
	22	P	59	C_b	8.0
	22	P	70	C_b	8.0
	22	P	74	C_b	8.0
	22	P	75	C_b	8.0
Group 3	23	P	49	C_b	8.0
	23	P	59	C_b	8.0
	23	P	70	C_b	8.0
	23	P	74	C_b	8.0
	23	P	75	C_b	8.0
Group 4	24	P	49	C_b	8.0
	24	P	59	C_b	8.0
	24	P	70	C_b	8.0
	24	P	74	C_b	8.0
	24	P	75	C_b	8.0

Group 5	25	P	49	C_b	8.0
	25	P	59	C_b	8.0
	25	P	70	C_b	8.0
	25	P	74	C_b	8.0
	25	P	75	C_b	8.0
Group 6	12	P	69	C_b	8.0
	13	P	69	C_b	8.0
	14	P	69	C_b	8.0
	15	P	69	C_b	8.0
	16	P	69	C_b	8.0
Collection 2 (domain 2)					
	DNA-resid	**DNA atom**	**Prot-resid**	**Protein atom**	**Distance**
Group 1	3	P	115	C_b	8.0
	3	P	125	C_b	8.0
	3	P	140	C_b	8.0
	3	P	141	C_b	8.0
Group 2	4	P	115	C_b	8.0
	4	P	125	C_b	8.0
	4	P	140	C_b	8.0
	4	P	141	C_b	8.0
Group 3	2	P	115	C_b	8.0
	2	P	125	C_b	8.0
	2	P	140	C_b	8.0

	2	P	141	C_b	8.0
Group 4	5	P	115	C_b	8.0
	5	P	125	C_b	8.0
	5	P	140	C_b	8.0
	5	P	141	C_b	8.0
Group 5	6	P	115	C_b	8.0
	6	P	125	C_b	8.0
	6	P	140	C_b	8.0
	6	P	141	C_b	8.0
Group 7	32	P	135	C_b	8.0
	33	P	135	C_b	8.0
	34	P	135	C_b	8.0
	35	P	135	C_b	8.0
	36	P	135	C_b	8.0

Table 3. Information used for the TATA-DNA binding. A collection with a single group, in which 20 of the restraints were enforced at any step of the simulation. Concretely, DNA Residue ID (DNA-resid), DNA atom, Protein Residue ID (Prot-Resid), Protein atom and distances (in Å) are indicated.

DNA-resid	DNA atom	Prot-resid	Protein atom	Distance (Å)
6	P	161	C_b	8.0
6	P	162	C_b	8.0
8	P	175	C_b	8.0
10	P	130	C_b	8.0
11	P	45	C_b	8.0
12	P	88	C_b	8.0
12	P	90	C_b	8.0
12	P	92	C_b	8.0
13	P	72	C_b	8.0
13	P	88	C_b	8.0
13	P	89	C_b	8.0
14	P	68	C_b	8.0
14	P	72	C_b	8.0
22	P	70	C_b	8.0
23	P	71	C_b	8.0
23	P	75	C_b	8.0
24	P	84	C_b	8.0
24	P	96	C_b	8.0
25	P	41	C_b	8.0
26	P	40	C_b	8.0
27	P	135	C_b	8.0
27	P	183	C_b	8.0
28	P	179	C_b	8.0
28	P	181	C_b	8.0

28	P	183	C_b	8.0
29	P	163	C_b	8.0
29	P	179	C_b	8.0
29	P	180	C_b	8.0
29	P	181	C_b	8.0

Table 4. Competitive binding simulation information for the P22 system. Each group contains information that can be satisfied with respect to one DNA sequence or the other. 12 of the 19 groups were enforced at any time in the simulation. Concretely, DNA Residue ID (DNA-resid), DNA atom, Protein Residue ID (Prot-Resid), Protein atom and distances (in Å) are indicated.

	DNA-resid	DNA atom	Prot-resid	Protein atom	Distance (Å)
Group 1	3	P	164	C_b	6.0
	43	P	164	C_b	6.0
Group 2	3	P	165	C_b	6.0
	43	P	165	C_b	6.0
Group 3	4	P	180	C_b	6.0
	44	P	180	C_b	6.0
Group 4	34	P	174	C_b	6.0
	74	P	174	C_b	6.0
Group 5	34	P	177	C_b	6.0
	74	P	177	C_b	6.0
Group 6	34	P	178	C_b	6.0
	74	P	178	C_b	6.0
Group 7	33	P	190	C_b	6.0
	73	P	190	C_b	6.0
Group 8	33	P	189	C_b	6.0
	73	P	189	C_b	6.0
Group 9	33	P	193	C_b	6.0

	73	P	193	C_b	6.0
Group 10	32	P	187	C_b	6.0
	72	P	187	C_b	6.0
Group 11	32	P	188	C_b	6.0
	72	P	188	C_b	6.0
Group 12	13	P	124	C_b	6.0
	53	P	124	C_b	6.0
Group 13	13	P	127	C_b	6.0
	53	P	127	C_b	6.0
Group 14	13	P	122	C_b	6.0
	53	P	122	C_b	6.0
Group 15	14	P	109	C_b	6.0
	54	P	109	C_b	6.0
Group 16	14	P	108	C_b	6.0
	54	P	108	C_b	6.0
Group 17	14	P	112	C_b	6.0
	54	P	112	C_b	6.0
Group 18	23	P	98	C_b	6.0
	63	P	98	C_b	6.0
Group 19	24	P	114	C_b	6.0

	64	P	114	C_b	6.0

Table 5. Competitive binding simulation information for the TATA binding protein. Each group contains information that can be satisfied with respect to one DNA sequence or the other. 6 of the 8 groups were enforced at any point in the simulation. Concretely, DNA Residue ID (DNA-resid), DNA atom, Protein Residue ID (Prot-Resid), Protein atom and distances (in Å) are indicated.

	DNA-resid	DNA atom	Prot-resid	Protein Atom	Distance (Å)
Group 1	28	OP1	213	OG	5.0
	60	OP1	213	OG	5.0
Group 2	23	OP1	102	NH2	5.0
	23	OP1	102	NH1	5.0
	55	OP1	102	NH2	5.0
	55	OP1	102	NH1	5.0
Group 3	24	OP1	109	NH2	5.0
	24	OP1	109	NE	5.0
	56	OP1	109	NH2	5.0
	56	OP1	109	NE	5.0
Group 4	24	OP1	116	OG1	5.0
	56	OP1	116	OG1	5.0
Group 5	10	OP1	222	NZ	5.0
	42	OP1	222	NZ	5.0
Group 6	9	O3'	222	NZ	5.0

	41	O3'	222	NZ	5.0
Group 7	12	OP1	124	NZ	5.0
	44	OP1	124	NZ	5.0
Group 8	8	OP1	200	NH1	5.0
	40	OP1	200	NH1	5.0