

## SUPPLEMENTAL MATERIAL 5

for the paper “Position dependencies in transcription factor binding sites”

A. Tomovic and E. J. Oakeley

**Table Sup5-1.** List of PDB and corresponding JASPAR ids used in the structural analysis. This is actually the intersection of the PDB and JASPAR databases. There are four additional, unused, structures (1IFI, 6PAX, 1FOS, and 1H89) which correspond to the JASPAR ids MA0050, MA0068, MA0099, and MA0100. They were not used because only position weight matrices, and not binding sites, are provided in JASPAR (September, 2006).

PDB ids	JASPAR ids	PDB id	JASPAR id
1EGW	MA0001	1YNW	MA0074
1H9D	MA0002	1FJL	MA0075
1AN4	MA0004	1BC8	MA0076
1EGW	MA0005	1PUE	MA0080
1R4I	MA0007	1K60	MA0081
1XBR	MA0009	1J46	MA0084
1K78	MA0014	1SKN	MA0089
1BY4	MA0017	1AN4	MA0093
1NWQ	MA0019	1B8I	MA0094
3HDD	MA0027	1UBD	MA0095
1DUX	MA0028	1GJI	MA0101
2HDC	MA0031	1A1G	MA0103
1AN2	MA0058	1NWQ	MA0102
1AWC	MA0062	1AN2	MA0104
1DSZ	MA0065	1SVC	MA0105
1B72	MA0070	1TSR	MA0106

**Table Sup5-2.** Structural analysis of co-crystal structures of TFs with DNA (**I** - at least one amino acid that interacted via hydrogen bonds or salt bridges with two binding positions, both of which were found to be dependent; **II** - at least one amino acid that interacted via hydrogen bonds or salt bridges with two independent binding site positions; **III** - at least one pair of dependent binding site positions with no apparent contact with the transcription factor; **IV** - at least one position of any pair of dependent positions had contact with transcription factors via hydrogen bonds or salt bridges).

Structural Characteristic	PDB id-JASPAR id
<b>I</b>	1EGW-MA0001, 1AN2-MA0058, 1FJL-MA0075, 1K6O-MA0081, 1AN2-MA0104, 1TSR-MA0106
<b>II</b>	1H9D-MA0002, 1AN4-MA0004, 1EGW-MA0005, 1R4I-MA0007, 1BY4-MA0017, 1NWQ-MA0019, 3HDD-MA0027, 2HDC-MA0031, 1AWC-MA0062, 1DSZ-MA0065, 1B72-MA0070, 1FJL-MA0075, 1BC8-MA0076, 1K6O-MA0081, 1J46-MA0084, 1AN4-MA0093, 1B8I-MA0094, 1GJI-MA0101, 1NWQ-MA0102, 1SVC-MA0105, 1TSR-MA0106
<b>III</b>	1EGW-MA0001, 1H9D-MA0002, 1EGW-MA0005, 1R4I-MA0007, 1XBR-MA0009, 1K78-MA0014, 1BY4-MA0017, 1NWQ-MA0019, 2HDC-MA0031, 1DSZ-MA0065, 1YNW-MA0074, 1K6O-MA0081, 1NWQ-MA0102, 1TSR-MA0106
<b>IV</b>	1EGW-MA0001, 1H9D-MA0002, 1EGW-MA0005, 1R4I-MA0007, 1K78-MA014, 1NWQ-MA0019, 1DUX-MA0028, 1AN2-MA0052, 1DSZ-MA0065, 1B72-MA0070, 1FJL-MA0085, 1PUE-MA0080, 1J4G-MA0084, 1K6O-MA0081, 1GJI-MA0101, 1NWQ-MA0102, 1A1G-MA0103, 1SVC-MA0105, 1TSR-MA0106

We found six pairs of PDB id-JASPAR id in which there was at least one amino acid that interacted via hydrogen bonds or salt bridges with two dependent binding positions. At least one amino acid interacted with two independent binding positions in 21 pairs. However, as discussed before, we anticipate that many of these may become dependent as additional binding site information becomes available. We found 14 pairs in which there was at least one pair of dependent binding site positions with no apparent contact with TF. There are 19 pairs in which at least one position of any pair of dependent positions had contact with TFs via hydrogen bonds or salt bridges.