In this weekly report I am summarizing recent studies on how DNA shape can be important for TF binding, while mentioning difficulties that may arise during interpretation of results. Following are examples for Pit1, Oct1, MEF2, P22, Phage 434 repressor, MSx1.

**Pit1**

In a 2002 study, a mutation in the promoter region of the GHRH receptor gene (A→C transversion at position -124) was found at the binding site of Pit1 in patient with growth retardation. The mutation changes the binding site which would normally be TGCAACAG to TGCCAAACAG, and changes GHRHR promoter activity.

![Diagram of DNA binding site](image)

However, the same region is also contacted from the major groove. The protein residue Thr45 makes contact with the mutated base in the major groove, and in the minor groove there is insertion of the Arg105

Salvatori et al., Decreased expression of the GHRH receptor gene due to a mutation in a Pit1 binding site, Molecular Endocrinology, 2002
Mutation in the consensus Oct binding site from ATGCAAAT to CTGCAGAT, which decreases promoter activity. However, the mutation is in 2 bps, which makes it difficult to separate the effects, besides the fact that there is interactions from both major and minor grooves. Sequence from crystal structure is 5’ GTATGCAAATAAGG ; bases in red are the ones mutated.

Reece-Hoyes et al., A consensus Oct1 binding site is required for the activity of the Xenopus Cdx4 promoter. http://dx.doi.org/10.1016/j.bbr.2011.03.031
Consensus MEF2 binding site is YTA(A/T)4TAR. The crystal structure has the following sequence 5’ CTATTTTAAA. Oishi et al., found that SNP rs3812852 located at position -1282bp (GTAAACTG to GTAA(GACTG) of the gene KLF5 is associated with hypertension.

A/A and A/G genotypes = higher risk
G/G genotype = lower risk

Regulatory region contains MEF2A binding site. A → G substitutions in vitro disrupts binding, and in the presence of AII substitution decreases promoter activity.

5’-GTGGTATATGTAA
A ACTGTCTAATG

5’-GTGGTATATGTAA
G ACTGTCTAATG

The A → G substitution contributes widen the minor groove in a region contacted by positively charged residues (R3, K4, R24, K31) according to the crystal structure from Wu et al.; however, just like in the other examples, part of the region responsible for the change in shape is also contacted in the major groove, but there is not as much specific base readout.

Another example in which the disruption of MEF2 binding site results in altered promoter activity is seen in Sofronescu et al. However, in that case multiple bases within the binding site are disrupted (from TATTTTTATTTTT to TAgTcTgAcTTTTT), and there is only evidence of a MEF2 element (doesn’t specify MEF2A, B, C...).

Oishi et al, Regulatory polymorphism in the transcription factor KLF5 at the MEF2 element alters response to angiotensin II and is associated with human hypertension. FASEB Journal, 2009
Wu et al., Structure of the MADS-box/MEF2 domain of the MEF2A bound to DNA and its implications for myocardin recruitment, JMB, 2010
Sofronescu et al., A Myocyte Enhancer Factor 2 (MEF2) Site Located in a Hypersensitive Region of the FGF16 Gene Locus Is Required for Preferential Promoter Activity in Neonatal Cardiac Myocytes, DNA and Cell Biology, 2008
For the P22, there is no base readout with the central 4 bps. In the 2013 paper, they suggest that affinity of a given sequence correlates with its ability to assume B’-DNA form. The following are the sequences studied and Kds for them. Note that at the moment we can only predict minor groove for 3 of them, since the others use I or U bases. Even though there are no arginines interacting with the minor groove, the presence of residues E44 and E48 are important for binding and loss of charge alters binding.

<table>
<thead>
<tr>
<th>Name</th>
<th>Seq</th>
<th>Kd</th>
</tr>
</thead>
<tbody>
<tr>
<td>9T</td>
<td>AGAT•ATCTTAAAT</td>
<td>1.0</td>
</tr>
<tr>
<td>9C</td>
<td>AGAC•GTCTTAAAT</td>
<td>10.6</td>
</tr>
<tr>
<td>9C/I</td>
<td>AGAC•ITCTTAAAT</td>
<td>0.7</td>
</tr>
<tr>
<td>9U</td>
<td>AGAU•ATCTTAAAT</td>
<td>0.8</td>
</tr>
<tr>
<td>9A</td>
<td>AGAA•ATCTTAAAT</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Harris et al Koudelka, Indirect readout of DNA sequence by p22 repressor: Roles of DNA and Protein Functional Groups in modulating DNA conformation, JMB, 2013
Phage 434 repressor

In Koudelka et al., 1987 we can see a representation of natural and synthetic operators (Fig a and b). Data on affinity of repressor for synthetic operators (Fig b) is available on this paper, but I haven’t found data on relative affinity for all sequences. These are mostly sequences with ACAAN6TTGT.

Sequences vary in non-contacted bases. From the natural operators, most of them have somehow a narrow groove at position 7/8/9 (looking at plot for minor groove width), but I need to find data that shows relative affinity for all of them together so they can be compared. Harisson et al., 1988 has relative Kds for natural Operators 1, 2 and 3.

If we look at sequences from synthetic operators, however, a heatmap of minor groove alone cannot explain relative affinity for the various sequences.

Koudelka et al., 1987. Effect of non-contacted bases on affinity of 434 repressor.

Harisson et al., 1988. Recognition of DNA sequences by the repressor bacteriophage 434

Rodgers and Harrison, The complex between the phage 434 repressor DNA-binding domain and operator site

Or3: structural difference between consensus and non-consensus half-sites. Structure, 1993
Msx1

In the case of Msx-1, there were a study with mutation on the consensus DNA binding site, from CTAATTG (WT) to CTCCGTG (MUT). Crystal structure is CACTAA TTGAAGG, so bases highlighted in red would be the mutated ones. Even though minor groove shape is dramatically affected, we cannot draw many conclusions from this paper because the mutated region does not affect transcriptional repression by Msx1.

Catron et al., Transcriptional repression by Msx-1 does not require homeodomain DNA-binding sites. Mol. Cell. Biol. 1995,

For next week, I will keep looking into effects of mutations in binding affinity/promoter activity while focusing more on GWAS and ways by which regulatory polymorphisms and changes in TFBSs could affect minor groove width and therefore give a phenotype, and how that could be tested.