Structure *14*, 555–565, March 2006 DOI 10.1016/j.str.2005.11.017

Principles of Protein-DNA Recognition Revealed in the Structural Analysis of Ndt80-MSE DNA Complexes

Jason S. Lamoureux[1] and J. N. Mark Glover[1,*]
[1] Department of Biochemistry
University of Alberta
Edmonton, Alberta T6G 2H7
Canada

Summary

The *Saccharomyces cerevisiae* transcription factor Ndt80 selectively binds a DNA consensus sequence (the middle sporulation element [MSE]) to activate gene expression after the successful completion of meiotic recombination. Here we report the X-ray crystal structures of Ndt80 bound to ten distinct MSE variants. Comparison of these structures with the structure of Ndt80 bound to a consensus MSE reveals structural principles that determine the DNA binding specificity of this transcription factor. The 5′ GC-rich end of the MSE contains distinct 5′-YpG-3′ steps that are recognized by arginine side chains through a combination of hydrogen bonding and cation-π interactions. The 3′ AT-rich region is recognized via minor groove contacts that sterically exclude the N_2 atom of GC base pairs. The conformation of the AT-rich region is fixed by interactions with the protein that favor recognition of poly(A)-poly(T) versus mixed AT sequences through an avoidance of major groove steric clashes at 5′-ApT-3′ steps.

Introduction

Key events during sporulation in *Saccharomyces cerevisiae* are triggered by waves of gene transcription that are driven by meiotic transcription factors. The genes whose transcription is activated during meiosis have been broadly classified as early, middle, mid-late, and late, based on the timing of their expression (Mitchell, 1994). The middle genes are activated in response to the successful completion of meiotic recombination, and are required for nuclear divisions and cell cycle progression (Hepworth et al., 1998; Xu et al., 1997). Of these middle genes, approximately 70% contain a conserved sequence in their promoter regions called the middle sporulation element (MSE), which is required for the activation of these genes during the middle phase of sporulation (Chu et al., 1998; Chu and Herskowitz, 1998; Hepworth et al., 1995, 1998).

Ndt80 is a transcription factor that has been shown to bind the MSE and activate the transcription of middle genes *in vivo* (Chu and Herskowitz, 1998; Xu et al., 1995). Ndt80 itself is classified as a middle sporulation gene, and its induction leads to positive autofeedback of its own transcription via its upstream MSE (Chu and Herskowitz, 1998; Pak and Segall, 2002). The Ndt80 protein is also subject to posttranslational modifications that appear to potentiate the transcriptional activation function of Ndt80 (Tung et al., 2000). Some MSE sequences have been shown to not only activate middle genes, but also to repress the transcription of these same genes during vegetative growth (Pierce et al., 1998, 2003). MSE-mediated repression requires Sum1, a DNA binding protein that binds a subset of the MSEs and recruits the Hst1 histone deacetylase to these genes, presumably to promote an inactive chromatin structure at these genes during vegetative growth and early sporulation (Xie et al., 1999). It has been suggested that the relative affinities of Ndt80 and Sum1 for different MSEs may in part control the differential timing and level of activation of middle sporulation genes (Pierce et al., 2003; Xie et al., 1999).

The structure of the DNA binding domain of Ndt80, both free and in complex with an MSE DNA, has been determined (Lamoureux et al., 2002; Montano et al., 2002). This work revealed that Ndt80 is a member of the Ig-fold family of transcription factors, which includes p53, NF-κB, STAT, AML-Runt, and the Rel subfamilies. All of these Ig-fold proteins bind DNA in a similar manner using loops and other features at one end of the β-sandwich (Bravo et al., 2001; Tahirov et al., 2001). However, Ndt80 seems to have the most extensive binding interface, utilizing additional secondary structure elements not seen in any other members of this family, allowing it to bind to its target DNA with high affinity as a monomer. Ndt80 also contains an N-terminal tail that appears to be novel in the family. In the absence of bound DNA, the tail is flexible, but becomes structured upon binding DNA, wrapping around the DNA in a manner that is essential for the specific recognition of the MSE.

The Ndt80-MSE structure has been refined to 1.40 Å, the highest resolution yet achieved for a transcription factor-DNA complex. The high-resolution structures (Lamoureux et al., 2002; Montano et al., 2002), together with studies of the effects of DNA and protein mutations on binding affinity (Fingerman et al., 2004; Lamoureux et al., 2002; Montano et al., 2002; Pierce et al., 2003), have shed light on how Ndt80 specifically recognizes the MSE. The MSE has been well defined based on statistical analyses of the promoters of middle genes (Chu et al., 1998; Chu and Herskowitz, 1998; Wang et al., 2005) and mutational analysis (Lamoureux et al., 2002; Pierce et al., 2003), which have yielded a 9 base pair MSE consensus: 5′-$g_1N_2C_3R_4C_5A_6A_7A_8W_9$-3′ (where lower case letters indicate semiconserved residues, R indicates a purine, N indicates any nucleotide, W indicates either a thymine or adenine) (Figure 1A). The MSE, like the targets of many Ig-fold transcription factors, contains a 5′ GC-rich portion, and a 3′ AT-rich region. The 5′ GC-rich region contains two highly conserved 5′-YpG-3′ dinucleotide steps at positions 3/4 and 5/6, which are recognized in a similar manner by two, distinct arginine residues (Arg111 and Arg177, respectively). Note that the YpG step is on the complementary strand to the MSE consensus sequence listed above. In each case, the guanine is specifically recognized by bidentate hydrogen bonds between the arginine guanidinium and the N_7 and O_6 atoms of the guanine (Figure 1B). The 5′-pyrimidine shifts toward

*Correspondence: mark.glover@ualberta.ca

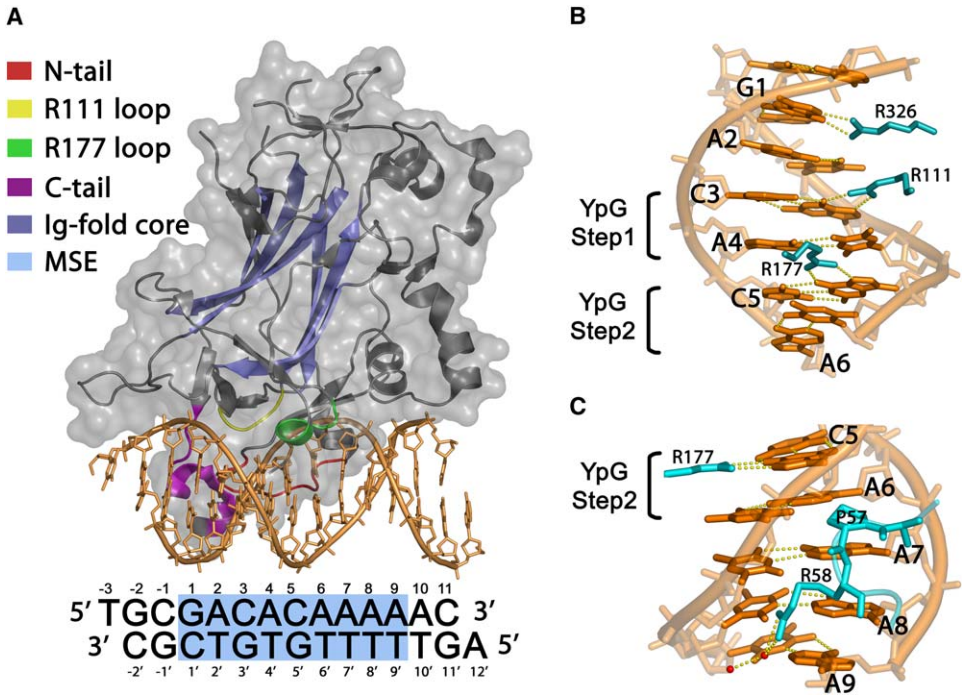

Figure 1. Recognition of the MSE DNA by Ndt80 in the Wild-Type Complex

(A) An overall diagram of the Ndt80-DNA complex with key DNA recognition loops highlighted. The DNA used in the wild-type structure is shown below with the MSE highlighted and numbered.

(B) Major groove view of the wild-type structure. Note the high degree of base unstacking of the 5′-pyrimidines in both 5′-YpG-3′ steps.

(C) Minor groove view of the wild-type structure. The steric constraints of residues Pro57 and Arg58 span four base pairs or the entire poly-A tract. The unstacking of the second YpG step is even more apparent in this view.

the major groove so that it no longer stacks with the 3′-guanine, but instead interacts with the arginine via cation-π interactions. Mutation of either YpG step, or the arginine residues that recognize them, lead to reduction DNA binding affinity, as well as a corresponding reduction in transcriptional activation in model reporter assays (Fingerman et al., 2004; Pierce et al., 2003). This mode of YpG recognition has been proposed to be utilized by several other classes of transcription factors, including the Zn-finger, bZIP, homeobox, HTH, and winged-helix families (Lamoureux et al., 2004). Recognition of the 3′ poly(A)-poly(T) segment largely involves the flexible N-terminal region of Ndt80. Pro57 and Arg58 are inserted into the minor groove, widening this region and making close contact with the base pairs in a way that would sterically occlude the N_2 group of guanine bases (Figure 1C). This explains the marked binding preferences for AT over GC pairs at positions 6–8. However, it was unclear why poly(A)-poly(T) is strongly selected over mixed AT sequences in this region of the MSE.

Many natural MSEs contain single base pair variations, which affect Ndt80 binding affinities and may be important in tuning their activities during sporulation. Here we report the high-resolution crystal structures of several Ndt80-DNA complexes containing MSE sequences with deviations from the consensus sequence that was used in the initial crystallization study (Table 1). Overall, the structures reveal that Ndt80 holds the DNA in a rigid conformation that is relatively resistant to structural variations in response to these changes. Analysis of these structures is consistent with a dominant role for recognition of the YpG steps in the 5′ part of the MSE, and provides an explanation for the selective recognition of poly(A)-poly(T) at the 3′ end over mixed AT sequences.

Results

The variant and mutant MSE sequences used in this study are summarized in Table 1 and a summary of the X-ray experiments are presented in Table 2. In most cases, we attempted to crystallize complexes containing a single base pair substitution that resulted in small but significant defects in binding affinity. The MSEs are classified into two groups: variants (specified with the prefix "v"), in which the MSE sequence is changed from the original or "wild-type" sequence but retains the minimal MSE consensus requirements; and mutants (specified with the prefix "m"), in which one of the conserved base pairs has been altered. In all cases the crystals contained the DNA binding domain of the Ndt80 protein from residues 1–340 and a 14-mer DNA with a single 5′ overhanging nucleotide. We attempted to crystallize complexes containing a total of 20 different mutant and variant MSEs. Of these, nine yielded crystals which diffracted X-rays to high resolution (between 1.56 and 2.0 Å). Of the 11 complexes that did not produce satisfactory crystals, 5 were single mutations at position 3 or 5 that were expected to significantly reduce the binding affinity, 3 were double mutations at semi- or nonconserved positions, resulting in variant MSEs, and 3 were single mutations of the remaining conserved positions that are expected to marginally decrease the binding affinity.

Table 1. Comparison of Previously Reported Binding and Activation Data and the Rmsd of the Variant Structures in Comparison to Wild-Type

Complex[a]	MSE sequence	Relative K_d[b]	Fold decrease in binding[c]	% Activation[d]	Rmsd (aligned to WT in Å)[e]	
					All atoms	DNA
WT (1MNN)	GAC ACA AAA	1	—	100	—	—
vG1A/A9T (2EUV)	AAC ACA AAT	—	—	—	0.65	0.40
vG1C (2ETW)	CAC ACA AAA	3.3	3.6	20	0.63	0.48
vA4G (2EUX)	GAC GCA AAA	3.1	5.9	20	0.52	0.33
mA4T (2EUW)	GAC TCA AAA	5.4	—	19	0.49	0.49
mC5T (2EUZ)	GAC ATA AAA	4.4	50	14	0.48	0.28
mA6T (2EVF)	GAC ACT AAA	4.0	8.3	18	0.50	0.31
mA7T (2EVG)	GAC ACA TAA	3.7	5.6	10	0.50	0.27
mA7G (2EVH)	GAC ACA GAA	—	7.1	37	0.51	0.23
mA8T (2EVI)	GAC ACA ATA	3.2	3.2	77	0.53	0.26
mA9C (2EVJ)	GAC ACA AAC	—	—	—	0.53	0.40

[a] The value in parenthesis is the PDB code.
[b] K_d of Ndt80 (1–340) for the indicated DNA substrate as referenced in the article by Lamoureux et al. (2002).
[c] Fold decrease of Ndt80 (1–409) bound in comparison to a wild-type substrate (Pierce et al., 2003). The reference reported percent bound values, whereas this column is 100%/percent bound values for easier comparison to the relative K_d values.
[d] Percent activation of the DNA sequence in a reporter assay as referenced by Pierce et al. (2003).
[e] Rmsd of the variant structure in comparison to wild-type using aligned atoms with the align function of PyMol (DeLano, 2002).

All of the crystals obtained are isomorphous with the wild-type, and have been refined to their respective resolution limits (see Experimental Procedures). The accuracy of these structures is sufficient to define not only the protein geometry, but also the DNA backbone torsion angles and the protein-DNA contacts, as well as networks of water molecules trapped at the protein-DNA interface that mediate recognition.

Variants

vG1A/A9T

vG1A/A9T was the only structure we obtained of a complex containing two substitutions. The substitutions are at semiconserved positions at the edges of the MSE that do not seem to play a major role in binding affinity, but nevertheless may be important in fine-tuning the relative affinities of a particular MSE for either Ndt80 or Sum1. The overall structure is very similar to the wild-type structure, with an rmsd for Cα atoms of 0.43 Å (Figure 2A). The most striking change in this variant occurs at position 1 of the MSE, where a guanine base is recognized by Arg326 in the wild-type structure. In this variant the guanine is replaced with an adenosine and the bidentate hydrogen bonding to the arginine cannot occur. Instead, the arginine side chain shifts to hydrogen bond with the backbone phosphate at the −1 position. The density in this position suggests that there is an alternate conformation of this arginine residue in which the

Table 2. Summary of X-Ray Experiments

Crystal data	1MNN	v1	G1C	mA4T	vA4G	mC5T	mA6T	mA7T	mA7G	mA8T	mA9C
A	70.13	69.23	70.3	69.99	69.96	70.28	70.19	70.43	69.53	69.71	69.63
B	78.81	79.25	78.92	78.34	79.13	78.84	78.75	78.78	78.44	78.55	79.04
C	161.39	160.88	161.63	161.48	161.83	161.65	161.54	162.67	161.33	161.78	161.68
Space group	C2221	C2221	C2221	C2221	C2221	C2221	C2221	C2221	C2221	C2221	C2221
Data collection											
Wavelength (Å)	0.980	1.072	1.072	1.072	1.072	1.072	1.072	1.072	1.009	1.009	1.009
Resolution (Å)	100–1.4	17–1.95	24–1.67	24–1.68	36–1.57	24–1.56	24–1.56	27–1.55	44–2.0	35–1.8	36–1.9
Unique reflections	88,384	35,074	52,485	50,924	62,985	64,128	63,947	65,883	30,725	41,465	35,986
Completeness	99.8 (97.8)	99.4 (100)	98.6 (99.0)	98.7 (99.5)	94.6 (72.2)	96.0 (86.9)	98.7 (93.4)	98.5 (90.8)	95.8 (81.3)	99.9 (100)	99.9 (100)
I/sigma	35 (3.0)	9.6 (1.4)	18.4 (3.8)	17.8 (1.6)	15.1 (3.0)	15.2 (2.1)	19.1 (3.1)	14.1 (2.5)	23.2 (2.4)	21.8 (2.5)	23.6 (2.73)
Rsym	0.051 (0.493)	0.045 (0.51)	0.043 (0.176)	0.067 (0.45)	0.051 (0.225)	0.051 (0.341)	0.045 (0.239)	0.053 (0.283)	0.067 (0.510)	0.090 (0.806)	0.074 (0.612)
Redundancy	38 (5.2)	4.1 (3.8)	4.7 (2.7)	8.9 (5.4)	4.3 (1.9)	4.6 (3.1)	5.6 (3.2)	4.9 (3.0)	3.4 (2.3)	7.5 (7.5)	3.8 (3.7)
Refinement											
Rcrys/Rfree	19.4/20.6	19.3/23.3	18.0/19.9	17.7/20.1	17.1/20.3	17.6/19.4	17.2/19.7	16.9/18.4	17.9/21.2	17.2/21.1	17.2/20.2
Rmsd bonds	0.011	0.011	0.010	0.010	0.010	0.010	0.010	0.009	0.008	0.011	0.008
Rmsd angles	1.52	1.45	1.38	1.39	1.45	1.42	1.41	1.40	1.33	1.45	1.27
Average B-factor, protein/DNA/ water (Å^2)	19.1/ 36.4/ 39.9	27.5/ 42.8/ 47.2	17.6/ 28.2/ 36.2	16.8/ 28.7/ 39.7	16.0/ 29.3/ 36.4	19.8/ 35.1/ 38.0	17.3/ 31.2/ 39.3	19.3/ 31.7/ 40.6	19.6/ 31.5/ 37.3	16.5/ 28.5/ 35.9	19.4/ 29.9/ 37.5
Ramachandran											
Favored/allowed	90.6/8.2	90.2/8.6	90.5/8.8	92.9/6.3	92.5/6.7	92.2/7.1	92.2/7.1	92.2/6.7	91.8/7.1	92.2/6.7	93.3/5.5
Generously allowed/ disallowed	0.8/0.4	0.8/0.4	0.4/0.4	0.4/0.4	0.4/0.4	0.4/0.4	0.4/0.4	0.8/0.4	0.8/0.4	0.8/0.4	0.8/0.4

Values in parentheses refer to the highest resolution shell.

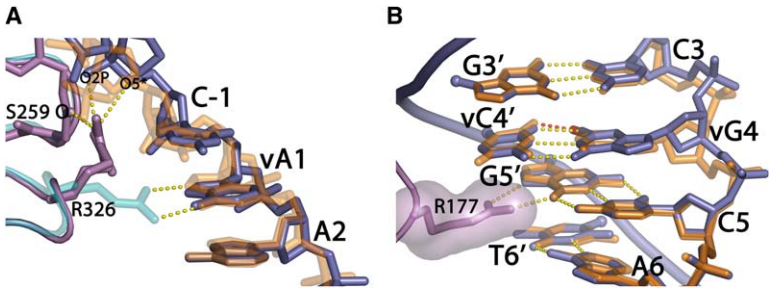

Figure 2. Structural Changes Seen in the Variant Complex Structures

In both panels, the wild-type structure is semitransparent and illustrated in orange and teal for DNA and protein, respectively. The variants are colored blue and purple for DNA and protein, respectively.

(A) Structural rearrangements seen in the vG1A/A9T complex. There is a large shift of Arg326 residue, which makes new contacts to the DNA backbone and also appears to have an alternate conformation.

(B) Structural rearrangements seen in the vA4G complex. The significant opening (~7°) of this base pair in the wild-type structure does not pose a problem for an A-T base pair; however, when replaced with a G-C pair, the guanine N_2 and cytosine O_2 distance would be too short (illustrated with a red hydrogen bond) if the bases are not rearranged.

Cδ-Nϵ angle is rotated 180°. In both conformations, there is a hydrogen bond to the carbonyl oxygen of Ser259 and weak hydrogen bonds to the phosphate. In response to the arginine reorientation, the DNA shifts such that the cytosine 5′ to the substituted adenine slides back ~1 Å to stack over the adenine, and the 5′ phosphate of the cytosine is repositioned to contact the arginine guanidinium. In the model, this shift is accomplished by a change in the backbone ϵ and ζ torsion angles from the common BI conformation (although not a proper BI), where ϵ and ζ are in the (t/g$^-$) range and $\epsilon - \zeta \approx -90$, to the less frequently observed BII conformation where ϵ and ζ are in the (g$^-$/t) range and $\epsilon - \zeta \approx +90$ at the −1 position (Prive et al., 1987). Interestingly the DNA appears to be poorly ordered in this structure in comparison to the others. In fact, this variant was the only one in which the density was poor enough to exclude modeling the nucleotides at the −2′ and −3 positions.

The other variation is at the opposite end of the MSE consensus, where the position-9 adenine is changed to a thymine. This change maintains the consensus for an MSE sequence. In the wild-type structure, Arg58 is packed into the minor groove in a conformation in which its aliphatic portion of the side chain (Cβ, Cγ, and Cδ) mimics Pro57 (Cβ, Cγ, and Cδ), and serves to exclude the N_2 of a guanine from the minor groove in the poly-A tract of the MSE. The guanidino group of this arginine coordinates a water molecule that appears to serve a similar role to exclude G-C base pairs at position 9 of the MSE. In this variant, the position of the DNA backbone and base pairs are extremely similar to the wild-type structure. The only minor variation occurs with a slight shift of the coordinated water accommodating the change of its hydrogen bonding partner from a nitrogen to an oxygen.

vG1C

The changes seen with the vG1C variant are very similar to those observed for vG1A/A9T (rmsd [Cα atoms] = 0.42 Å). Arg326 cannot hydrogen bond the substituted cytosine at the +1 position, and therefore swings out to contact the DNA backbone, as is seen in vG1A/A9T. However, there are two differences between these two variants. First, this arginine does not appear to have alternate conformations. Second, the shift of the DNA backbone is not accomplished with a switch from the BI to BII conformation, but merely a displacement of the DNA with maintenance of torsion angles closer to the wild-type structure. This variant also lacks interpretable density for the 5′ overhanging thymine; however, in contrast to vG1A/A9T, the −2′ base pair can be modeled in this variant. A caveat for both vG1C and vG1A/A9T mutants is that, due to the relatively poor electron density around position 1, torsion angles can not be accurately determined; however, the shift from BI to BII can be reasonably inferred, as it involves a characteristic shift in the relative position of the $O_{3'}$ atom.

vA4G

The final complex of a variant MSE bound to Ndt80 was obtained by replacing the position-4 AT pair found in the wild-type structure with a GC, maintaining the requirement for a purine at position 4. Remarkably, despite maintaining a consensus MSE sequence, this variant had the largest DNA backbone shift seen in all of the complexes investigated in this study (Figure 2B). The driving force for this structural rearrangement probably lies in the fact that the TA pair in the wild-type structure is significantly opened toward the major groove by about 7° relative to the mutant. While this opening does not dramatically distort the two hydrogen bonds of the AT pair, if one simply replaces the AT pair with a GC in this opened conformation, the guanine N_2-cytosine O_2 distance is too short for a stable hydrogen bond (2.3 Å). The cytosine in the variant also shifts slightly (~0.3 Å) into the major groove, probably to increase stacking interactions with Arg111, but exerting further pressure on the guanine partner. As a result, the guanine of the opposite strand shifts out 1.4 Å into the major groove to maintain good hydrogen bonding geometry, facilitated by a shift to the BII conformation in this nucleotide. The fact that the guanine nucleotide is not contacted by the protein probably allows this movement and, as a result, there is only a small loss of binding affinity associated with this mutation.

Mutants

mA4T and mA6T

Each of these mutations destroys the MSE consensus by mutating the 5′-pyrimidine in one of the two critical 5′-YpG-3′ steps to an adenine, resulting in 5′-ApG-3′ steps. What is most remarkable about these mutations is the lack of change in the structures (Figure 3). Despite disrupting the consensus MSE, the conformation of both the DNA and protein is almost entirely unaltered. Even the hydrogen bonding in the minor groove is maintained because the positions of hydrogen bond acceptors do not change. Only in the major groove does the hydration shell vary, as would be expected with a change in the

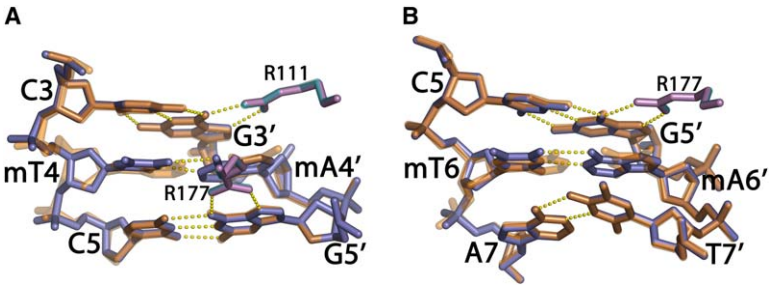

Figure 3. Mutation of the Pyrimidines in the 5′-YpG-3′ Steps

In both panels, the wild-type structure is semitransparent and illustrated in orange and teal for DNA and protein, respectively. The mutants are colored blue and purple for DNA and protein, respectively.

(A) mA4T complex.

(B) mA6T complex.

In both cases, despite the relatively large decrease in binding affinity, there are surprisingly few structural changes. The expected relaxation of unstacking when the YpG step is mutated was not observed; instead, the DNA appears to be anchored in place.

positions of the hydrogen bonding partners. The DNA backbone is held rigidly in place through extensive contacts with Ndt80, such that the substituted 5′-adenosine is held in the BII conformation shifted into the major groove. In spite of the fact that the N_9 of the adenine in the mutants, and the N_1 of the corresponding thymines in the wild-type structure, are in virtually the same position, the reduced size of the purine imidazole, compared to the thymine, results in significantly less van der Waals contacts between the adenine and the arginine.

The conformational energy of the dinucleotide steps in each of these structures can be estimated using helical parameters. The energy values are expressed in terms of $k_BT/2$ relative to the mean values for free B-DNA (Olson et al., 1998). For mA4T, the energy of the mutated ApG step is 67, and for mA6T it is 64, whereas the corresponding TpG steps in the wild-type structure are 44 and 27, respectively. It is important to note that these two TpG steps are already the highest energy steps in the wild-type structure; indeed, the average conformational energy for dinucleotide steps in the wild-type structure is 14. The reduced van der Waals contact, together with the additional cost of unstacking the more rigid dipurine step, largely explains the small but biologically significant loss of binding affinity (~2- to 3-fold) for these sequences compared to the consensus MSE.

mC5T

Our previous studies demonstrated that the conserved CG base pairs at positions 3 and 5 are the most sensitive to mutation. Although we attempted to crystallize all possible substitutions at these positions, we were only able to crystallize mC5T. This mutation is the best tolerated of the three possible mutations at this position, and results in a 3-fold decrease in affinity in comparison to the wild-type MSE (Lamoureux et al., 2002). This structure has the lowest rmsd (all aligned atoms) of all the variants and mutants (rmsd = 0.48 Å). The reason mutations are poorly tolerated in this position is that the guanine at position 5′ along with the thymine at position 6′ comprise the central 5′-YpG-3′ step of the MSE consensus. This step is recognized by Arg177 in the major groove and Pro57 in the minor groove as well as by extensive backbone phosphate and sugar contacts (Figure 1C). This region of the DNA is almost completely encompassed by Ndt80 and accounts for a great number of the protein-DNA contacts. The substitution of the CG by the TA base pair at this position makes it impossible for Arg177 to make bidentate hydrogen bonds to the guanine base. Instead, Arg177 rotates ~180° about χ2, displacing two water molecules in the process (Figure 4). The two displaced waters are part of a cluster of six well-ordered waters that mediate interactions between the protein and base pairs 4 and 5. The guanidinium group of Arg177 in the C5T structure makes hydrogen bonds that replace some but not all of those made by the displaced water molecules. As a result of this weakened hydrogen bonding network, the Arg177 guanidinium group is not held in place as tightly as in the wild-type structure, resulting in higher B factors for the guanidinium atoms (~30 Å^3) compared to the wild-type structure (~18 Å^3).

mA7T

In the wild-type structure, the poly-A tract is contacted predominantly through the minor groove and the backbone, and beyond position 6 of the MSE there are no direct side-chain contacts with the major groove at all (Figure 1C). As mentioned above, G-C base pairs are excluded in this region by the steric clashes that would occur in the minor groove by a 2-amino group of a guanine base, but it is not clear why an A-T to T-A base pair substitution is disfavored. This mutation introduces no new steric clashes with the protein, yet the affinity is reduced by approximately 3- to 5-fold (Figure 5A) (Lamoureux et al., 2002; Pierce et al., 2003). The position of the backbone and bases at the substituted thymine at position 7 remains quite similar to that of the wild-type structure despite a subtle shift from BI towards a more BII-like conformation (from $\epsilon - \zeta = -28$ in wild-type to +12 in mA7T). At position 6, there is the opposite shift from the rarer BII to the typical BI conformation, which repositions the adenosine base towards the minor groove by approximately 0.6 Å. Previous work has noted that the BII conformation is often associated with an unstacking of the base relative to the 3′ base (Prive et al., 1987). The change from BII to BI at position 6 and the additional BII character at position 7 seen in this mutant serve to increase the distance between the adenosine at position 6 and the introduced mutant adenosine on the opposite strand at position 7, yet the $A_6(N_6)$-$A_{7'}(N_6)$ distance is still 3.0 Å, indicating that these two atoms clash despite the rearrangements of the DNA. This clash, in addition to the rearrangements that are required to minimize it, are quite likely responsible for the 3- to 5-fold decrease in affinity of Ndt80 for this substrate.

mA8T

This mutant is similar to mA7T both in the type of mutation and its context within the MSE (Figure 5B). In addition, similar structural changes are observed in this

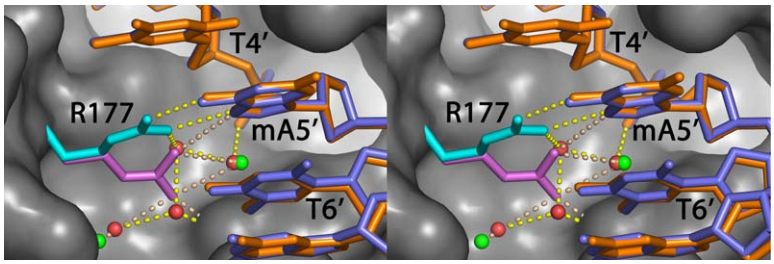

Figure 4. Stereoview of the Changes in the mC5T Structure

The mutation of the guanine recognized by Arg177 to an adenosine forces the side chain to reorient, displacing two water molecules. The wild-type structure is semitransparent, and illustrated in orange and teal for DNA and protein, respectively; waters are red and hydrogen bonds are yellow. The mutant structure is colored blue and purple for DNA and protein, respectively. The waters in the mutant are colored green and the hydrogen bonds are beige. The surface of the protein, excluding the side chain for Arg177, from the mutant structure is rendered in grey and is essentially identical to that of the wild-type structure.

mutant, although to a lesser degree. As is seen in the mA7T mutant, the base pairs of mA8T are in nearly identical positions as the wild-type structure. It is the adenosine at position 7 that, once again, shifts toward the minor groove by approximately 0.3 Å. This shift is accomplished by several changes in the backbone torsion angles, but, unlike mA7T, there is not a clear shift from BII to BI, as the wild-type structure is not in a BII conformation at this position. However, the adjustment at position 7 makes the mutant more "BI like" as the ϵ-ζ 5′ to the mutated adenine is −65° versus −28° seen in the wild-type structure. The additional BI character is likely due to a steric clash that would otherwise occur between the 5-methyl group of the mutated thymine and its phosphate. The thymine paired with the shifted adenine at position 7, however, remains in an identical position to that seen in the wild-type structure, due to the extensive backbone contacts with this strand of the DNA, effectively locking the DNA in place. Again it appears that these rearrangements serve to increase the distance between the adjacent adenines on opposite strands, resulting in an N_6-N_6 distance of 3.1 Å across the major groove.

mA7G

The structure of the wild-type complex suggested that GC base pairs would be disfavored at this position due to steric clash between the 2-amino group of the guanine base and Arg58 in the minor groove. The structure of the mutant reveals that, while the substituted cytosine adopts the same conformation as the thymine in the wild-type structure, the guanine base shifts to reduce its steric clash with Arg58 (Figure 6A). However, the magnitude of this shift is modest—only about 0.4 Å toward the major groove. This does not completely relieve the steric clash, as the guanine N_2-Arg58 Cγ distance is still a rather short 3.2 Å, in contrast to the 4.1 Å adenine C2-Arg58 Cγ distance observed in the wild-type structure. A change from the BI to BII conformation at position 7 assists the shift of the guanine base toward the major groove, and the concurrent change from BII to BI at position 6 serves to bring the rest of the DNA back into register with the wild-type structure.

The protein does not seemingly have any noticeable changes; in particular, Arg58 is in a nearly identical conformation as in the wild-type structure. It is interesting that it does not change conformation to avoid the close proximity to the 2-amino group of the mutant guanine, as most of the volume around this side chain contains water molecules that one might think could be easily displaced. This is not the case, and either the contacts made by this arginine hold it securely in place and/or the waters around this arginine are integral components of the structure. The B-factors of nearly all of these waters are <30Å, and their positions are conserved, consistent with the view that they are indeed critical to the structure, and hence the recognition of the MSE.

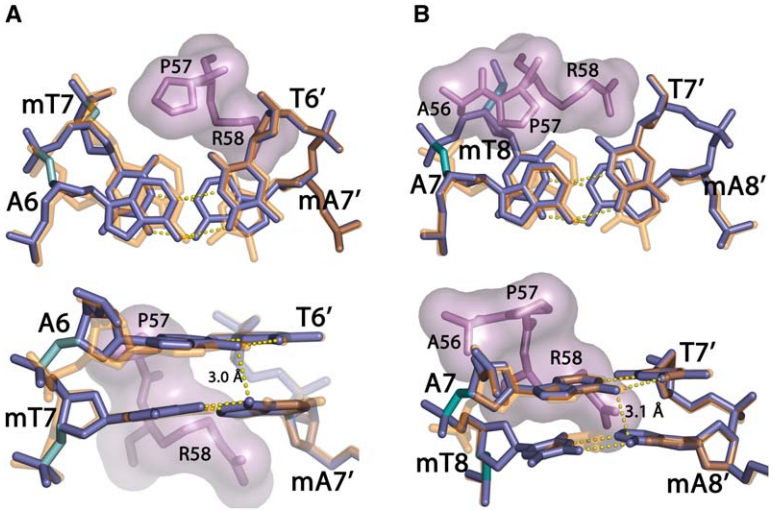

Figure 5. Changes in the Poly-A Tract Introduced with Alternating A-T Mutations

The mutants are colored blue and purple for DNA and protein, respectively with the ϵ and ζ bonds that define the BI and BII conformations highlighted in cyan, superimposed on the DNA from the wild-type structure (semitransparent orange). The top portion of each panel is a view down the helical axis, and the bottom portion is a perpendicular view from the major groove.

(A) Structural rearrangements of the mA7T mutant. The change from a BII to BI conformation at position 6 and the concurrent shift from BI to BII at position 7 serves to maximize the cross-strand N_6 distance.

(B) Structural rearrangements of the mA8T mutant. This mutant sees similar changes as mA7T, but to a lesser extent. There is also a potential clash between the position eight 5-methyl group and its phosphate oxygen.

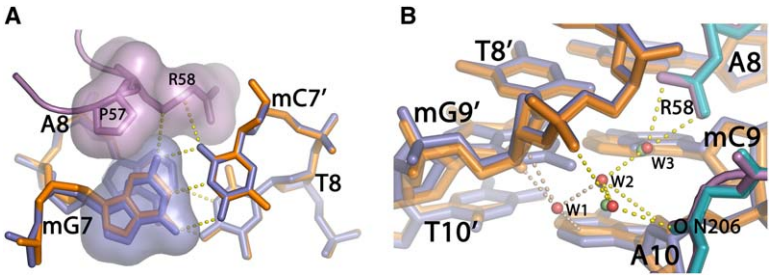

Figure 6. Other Changes in the Poly-A Tract

In both panels, the wild-type structure is semitransparent and illustrated in orange and teal for DNA and protein, respectively. The mutants are colored blue and purple for DNA and protein, respectively.

(A) Structural consequences of mA7G. Mutation to a guanine introduces an N_2 amino group and, consequently, a clash with the Arg58 side chain.

(B) Structural consequences of mA9C. The red spheres are the waters found in the wild-type structure, while the green ones are those of the mutant structure. Hydrogen bonds in yellow are those found in both the wild-type and mutant structures, while those in beige are wild-type hydrogen bonds that are lost in the mutant.

mA9C

This mutant structure is interesting in that it provides a possible explanation for the specificity at this position that was lacking in the original paper (Figure 6B). The only contact between this position and the protein is made through backbone contacts and water-mediated hydrogen bonds in the minor groove. In the original structure, we did not see why G-C base pairs were excluded from this position. It appeared that the 2-amino group of a guanine base could simply displace a water molecule in the hydrogen bonding network of the minor groove. This mutant structure demonstrates that the consequences of such a mutation may not be as trivial. In the wild-type structure, three of the waters that make up this hydrogen bonding network (W1, W2, and W3) are tightly hydrogen bonded to their maximum number of hydrogen bonding partners: two donors and two acceptors (Figure 6B). Nearly all of these hydrogen bonds are relatively strong, with the majority having a distance of less than 3.0 Å. The introduction of a guanine base displaces a single water molecule (W1), but the N_2 amino group is too far away to contact the hydrogen bonding partners of this displaced water. In effect, this mutation eliminates all four of the hydrogen bonds of the water it displaced. The whole network of waters in the minor groove seems to become destabilized if one uses B-factors as an indicator of relative stability. Despite the fact that the average B-factor for this mutant is lower than that of the wild-type structure, the remaining two waters in the mutant structure (W2 and W3) are both 10 $Å^3$ higher than that seen in the wild-type structure, indicating that these waters are destabilized. Interestingly, Arg58 has essentially identical B-factors in both the wild-type and mutant structure, indicating the difference seen in the B-factors of these waters is not simply a general localized disorder. If this explanation is indeed how Ndt80 recognizes position 9, then we would expect it to be an extremely weak preference for an A or T. In fact, two recent papers call into question the importance of this position for defining an MSE (Pierce et al., 2003; Wang et al., 2005).

Discussion

This large set of mutant and variant complex structures provides a basis for understanding the fundamental and subtle mechanisms of recognition utilized by Ndt80 in order to bind to its appropriate MSE target sequences. There is good agreement between the structural data and the mutational studies, both of the MSE and protein residues. Mutational studies of Ndt80 highlight the importance of Arg111, Arg177, Pro57, and Arg58 in MSE recognition and affinity (Fingerman et al., 2004; Montano et al., 2002). In addition, there are several residues that these studies highlight that cannot be directly rationalized through structure analysis alone. While the idea of a simple universal code that could be used to predict DNA-binding preferences of a given protein a priori seems unrealistic, perhaps the principles of recognition shown here might be applicable across transcription factor families and provide more insight into protein-DNA interactions as a whole. One of the most striking observations that can be made for these structures is their very high similarity to the wild-type structure. The structure with the largest rmsd was vG1A/A9T, with an rmsd of 0.66 Å over all aligned atoms (~2800 atoms) or 0.43 Å over the Cα atoms (289 atoms). This is a very close alignment when compared to the rmsd with the unbound structures of Ndt80, which vary from 0.83 Å (240 atoms-1MN4) to 2.67 Å (252 atoms-1M6U-A). Although this is not altogether unexpected, due to the crystal packing constraints and changes induced by DNA binding, the extent to which the complex resists large structural changes is remarkable. What is unexpected is the trend that these mutant structures display. As the mutations and variations are made closer to the middle of the MSE, where changes have the greatest effect on specificity and affinity, the rmsd actually decreases. One might expect that mutations that have the most detrimental effect on binding induce larger changes in the complex structure, but the opposite effect is seen. The most significant changes in these structures as a whole tend to be on the nonprime strand, where there are less protein contacts and rearrangements of the side chains that contact mutated base positions. In all cases, the structural changes are localized around the sequence changes, and do not propagate more than 2 base pairs from the mutation.

GC-Rich Region Recognition

The MSE can be crudely separated into two regions: a GC-rich region at one end of the consensus, and

a poly-A region at the other end. The GC-rich region is contacted in the major groove by Arg326, Arg111, and Arg177, residues that each recognizes distinct YpG steps in the MSE. The TpG steps at positions 3/4 and 5/6 are well defined in the strict MSE consensus (Chu and Herskowitz, 1998; Hepworth et al., 1995; Ozsarac et al., 1997). The 5′-pyrimidine of the CpG at −1/1 in our structure is only weakly conserved in the newly proposed MSE (5′-YGNCACAAAA-3′) (Pierce et al., 2003). Five of our complexes test the consequences of disrupting these steps by either mutating the 5′-pyrimidine or the 3′-guanine. Mutation of the 3′-guanine (seen in complexes vG1A/A9T, vG1C, and mC5T) is accommodated by rotation of the arginine away from the mutated base. In the case of the mutations at the position 1 guanine, the 5′-pyrimidine no longer stacks on the arginine (Arg326) and shifts back towards the DNA duplex axis to stack on the 3′-guanine. This shift is in agreement with our proposal that the unstacked conformation seen in the wild-type complex is energetically unfavorable in the absence of contacts with Ndt80. However, it is also likely that the new contact between Arg326 and the 5′-phosphate of the shifted pyrimidine also stabilizes this conformational change.

In contrast, mutation of any of the base pairs in the conserved YpG steps at either position 3/4 or 5/6 is not associated with a restacking of the DNA bases. One reason for this is that the region of the DNA from positions 3 to 6 has a large number of contacts to the protein in both the major and minor grooves as well as the DNA backbone. These extensive interactions likely anchor the DNA in its conformation with little regard for the changes made to the sequence. The energetic contributions of these contacts are high enough to compensate for the extra energy required to force the mutated dinucleotide step into the unstacked conformation, although this likely reduces the overall affinity of Ndt80 for these DNA substrates. It is interesting that the majority of these contacts are to only one strand of the DNA, the one that contains the YpG step, rather than its complementary strand.

The YpG step at positions 3 and 4 (gNCRCAAAA/T) contains an ambiguous YpG step originally thought to have no preference for a particular pyrimidine. However, recent work suggests that thymines are preferred over cytosine (Pierce et al., 2003). We have crystallized both the TpG (wild-type) and CpG (vA4G) variations, and the differences in these structures suggest an explanation for this discrepancy. First, as described above, the standard base pairing geometry of the 5′ CG base pair in the vA4G structure imposes a shift in the geometry of the backbone of the introduced guanine to the less stable BII conformation. Second, in the wild-type structure, the aliphatic portion of Arg177 forms a hydrophobic half pocket that cradles the 5-methyl group of the thymine (see Figure 2B). We have shown that this methyl group is important for binding in the analogous TpG step in the fifth and sixth positions of the MSE with binding studies that show an approximately 2-fold reduction in binding affinity when this base is changed to a uracil (Lamoureux et al., 2002). When this thymine is mutated to a cytosine, this base shifts slightly toward the major groove, possibly in an attempt to fill the hydrophobic "hole" left where the 5′-methyl of the thymine would be. This shift further increases the unstacking of this step and likely increases the energy required to achieve this conformation, thus lowering the overall affinity.

Poly-A Tract Recognition

The second region of the MSE is the poly-A tract, which is recognized primarily through minor groove and backbone interactions. Because there are no major groove contacts to the bases past position 6 of the MSE, it was difficult to fully rationalize the selection of a poly-A tract in terms of direct readout. The minor groove interactions of Pro57 and Arg58 serve to exclude G-C or C-G base pairs due to the steric clash that would occur with the 2-amino group of guanine, however the preference poly-A/poly-T over mixed A-T sequences was difficult to explain based on the structure of the wild-type complex alone. In the original paper of the wild-type structure, we predicted that preference for a poly-A tract might lie in the additional entropic cost of binding a flexible alternating A-T tract over a more rigid poly-A region that is preset in a conformation more suitable for Ndt80 binding. Several of the mutants crystallized here involve the poly-A tract, and suggest an alternate structural mechanism that may account for this specificity.

At this point it is useful to consider the structure of poly-A tract DNA and some of its hallmarks. Poly-A DNA structures in the unbound state typically have a high propeller twist that enhances intrastrand base stacking as well as promoting interstrand hydrogen bonding between adenine N_6 and thymine O_4 of adjacent base pairs. Additionally, poly-A (as well as mixed A-T) regions tend to have an especially narrow minor groove of approximately 9.5 Å versus the 12 Å seen in fiber B-DNA (Nelson et al., 1987; Yoon et al., 1988). The poly-A region of the MSE DNA in the wild-type complex does not exhibit high propeller twist nor is the minor groove narrowed, due to the insertion of the proline and arginine side-chains. In fact, at its widest point, near the 5′ end, the minor groove of the poly-A tract is 14.1 Å. The widened minor groove brings the adenine N_6 and thymine O_4 of adjacent base pairs closer together, such that high propeller twist is no longer necessary to establish hydrogen bonding between the adjacent base pairs. In fact, a high propeller twist in this region would cause the hydrogen bonding pair to clash.

To address the preference for poly-A over alternating A-T, there are two mutants, mA7T and mA8T, which introduce an alternating A-T region and allow us to see the effect of this change. As is seen with many of these mutants, the DNA backbone has some small changes while by and large retaining the wild-type structure. In both of these structures, the largest structural changes are not observed in the mutated nucleotides, but rather in the nucleotides 5′ to the mutation on the otherwise poly-A strand. Note it is the poly-A strand that has fewer protein contacts and is expected to be more flexible than the poly-T strand. The adenosine 5′ to the mutated base shifts 0.7 Å toward the minor groove in the mA7T structure and 0.4 Å in mA8T structure. This shift helps to reduce steric repulsions between the two N_6 atoms of adenines of adjacent base pairs at the introduced 5′-ApT-3′ step. In spite of these movements, the N_6 atoms are still close enough to significantly repel one another, 3.0 and 3.1 Å apart in the mA7T and mA8T

structures, respectively, which likely accounts for the lower binding affinity of these mutant MSEs (Figure 5). This effect is greatest at the 5′ end of the poly-A tract (position 6/7), where the minor groove is the widest and diminishes toward the 3′ end, so that at position 9, where the minor groove width is similar to unbound poly-A DNA (9.5 Å at position 9), there is little or no preference for A-T over T-A base pairs.

In addition to the N_6 clash of adjacent adenines in mA8T, the mutated thymine also introduces a potential clash between its 5-methyl group and its 5′-phosphate. This clash is alleviated by the backbone adopting a more "BI-like" conformation that shifts the phosphate group away from the 5-methyl, which has an additional effect of moving the 5′ adenine in the right direction to minimize the N_6-N_6 clash.

Implications for Modeling Meiotic Transcriptional Activation in *S. cerevisiae*

The structures presented here provide a detailed view of how Ndt80 binds a number of different DNA targets with subtle changes that modulate binding affinities within an order of magnitude. Classic experiments on the λ phage repressor/cro system have provided the best-known example of how differences in affinity of transcription factors for different DNA targets can drive a developmental program (in this case, the switch from lysogenic to lytic growth) in response to changing levels of transcription factor concentrations (Ptashne, 1986). It is tempting to speculate that a similar mechanism may regulate the precise timing of gene expression during progression through meiosis, dependent on the relative binding affinities of Ndt80 and Sum1 for key regulatory MSE elements (Pierce et al., 2003; Xie et al., 1999). For example, the activation of certain genes very soon after the recombination checkpoint might be explained by a relatively high affinity of Ndt80 and/or a relatively low affinity for Sum1 in key regulatory MSEs. Conversely, activation of other genes could be delayed until later in the developmental program, when Ndt80 protein levels are higher and Sum1 is lower, by the utilization of MSE elements, which have a correspondingly lower affinity for Ndt80 and/or higher affinity for Sum1. This differential timing of middle genes depends on whether these genes have a Sum1 binding site, an Ndt80 binding site, or both. In addition, the relative affinities of these binding sites can further influence the timing and result in activation of middle genes over the entire time course of middle sporulation. In fact, the four waves of gene expression originally used to describe sporulation were expanded to seven in a DNA microarray study that focused on sporulation (Chu et al., 1998). This work indicates that, as the transcription program is investigated on a finer scale, more temporal patterns will emerge.

Ndt80 and Sum1 have overlapping although distinct MSE binding site requirements and the binding of one of these proteins to its MSE is mutually exclusive (Pierce et al., 2003). These properties, combined with the fact that Ndt80 auto-induces its own expression in a positive feedback loop, have been used in computer models to generate a network with an extremely sharp expression profile (Wang et al., 2005). It has been speculated that both autofeedback and activator/repressor competition may be general features of processes that require sharp temporal and/or spatial gene expression, such as sporulation in yeast or developmental/differentiation pathways of higher eukaryotes (Wang et al., 2005). If this is the case, then the Ndt80/Sum1/MSE system could provide an ideal model to study this type of regulatory network.

The sharp expression profile of Sum1/Ndt80 dual-regulated middle genes, in addition to the potential differential timing of middle genes, forms an interesting and complicated expression system. In vivo, this system is further complicated by controls at the level of protein synthesis and degradation. Furthermore there are post-translational modifications that change the activity of these transcription factors, possibly in response to recombination checkpoints. In addition, there may be other transcription factors that enhance or antagonize the activity of Ndt80 or Sum1. If we hope to understand this system as a whole in vivo, a first step is to understand the basic principles of the DNA-protein recognition.

Conclusions

We have demonstrated the importance of protein contacts to the DNA backbone in molding the conformation of the DNA, even when the MSE is not conserved. This highlights the contribution of indirect readout on MSE recognition. The critical recognition of YpG steps in the MSE is mediated in part by indirect readout of the intrinsic flexibility of the pyrimidine-purine step, as well as the by direct readout of the base pair surfaces exposed in the major groove through hydrogen bonding, cation-π, and van der Waals interactions. Indirect readout seems to be implicated in poly-A tract recognition as well. It appears that Ndt80 does not specifically recognize the conformation of the poly-A tract in its unbound state; rather, this region of the MSE is recognized by its ability to adopt a conformation induced by Ndt80 binding while simultaneously avoiding steric clashes within the DNA and with the protein. The adjustments in the DNA are often accommodated by subtle shifts between BI and BII backbone conformations, allowing flexure of the double helix to adapt to its protein partner.

Experimental Procedures

Protein expression and purification, as well as DNA purification procedures, have been previously described (Lamoureux et al., 2002).

Crystallization and Data Collection

Ndt80 (1–340)-MSE complexes were prepared to a protein concentration ranging from 10 to 20 mg/ml and a protein:DNA ratio of 1:1. Crystals were grown at room temperature (20ºC) using the hanging drop method in conjunction with streak seeding using wild-type crystals. Variant G1A/A9T was used in the initial crystallization trials of Ndt80-DNA complexes, but did not yield crystals until it was streak seeded using wild-type Ndt80-DNA complex crystals. In most cases, once mutant crystals were obtained, further optimization was done using the mutant crystals for seeds in the streak seeding procedure. The reservoir solution contained 25%–35% PEG 400, 50 mM bis-tris-propane (pH 7.0), 100 mM NaCl, 50 mM $CaCl_2$, and 1–5 mM DTT; 2 μl of complex was mixed with 2 μl of reservoir to form the drop. Drops were streaked either immediately or after a day of equilibration. Streaking was done using a horse hair dipped in streaking solution and rinsed twice in the reservoir. The streak seeding solution was prepared by diluting a drop that contained small crystals ~100-fold using the reservoir solution. Crystals grew to a maximum size of ~100–400 μm in 1–2 weeks, and were harvested

and frozen in reservoir solution containing up to 10% glycerol where necessary. Data was collected at SBC-CAT (BL 19-ID) at the Advanced Photon Source (APS) and at the Advanced Light Source (ALS) (BL 8.3.1). All crystals belong to space group $C222_1$ (a = 70 Å, b = 79 Å, c = 161 Å ± 2%), with one complex in the asymmetric unit, and are isomorphous with the wild-type crystals (Table 2). All data from APS was processed with HKL2000 (Otwinowski and Minor, 1997), and all data from ALS was processed with mosflm and scala in the CCP4 program suite (CCP4, 1994).

Structure Determination, Refinement, and Analysis

The variant and mutant complexes were built using the wild-type Ndt80-MSE complex as the starting point. All waters were removed from the original pdb as well as the bases of mutated positions and the bases immediately adjacent to the mutated position(s). The remaining protein-DNA model was used to phase $2F_o$-F_c and F_o-F_c maps using the mutant diffraction amplitudes. Manual modeling was done with O (Jones et al., 1991) and, alternatively, PyMol (DeLano, 2002). Refinement and the addition of waters were carried out using REFMAC (CCP4, 1994), and protein geometry was analyzed with PROCHECK (Laskowski et al., 1993). The final models of all of the mutants are similar to wild-type containing residues 33–139, 146–286, 294–335 of Ndt80, most if not all of the DNA, and between 280 and 350 water molecules. The atomic coordinates of the NDT80-MSE variants and mutants have been deposited in the Protein Data Bank (PDB ID: vG1A/A9T [2EUV], vG1C [2ETW], mA4T [2EUW], vA4G [2EUX], mC5T [2EUZ], mA6T [2EVF], mA7T [2EVG], mA7G [2EVH], mA8T [2EVI], mA9C [2EVJ]). Structures were aligned using the align function of PyMol with default settings, which include two iterative cycles of outlier rejection. DNA helical parameters and torsion angles of the structures were calculated using 3DNA (Lu et al., 2000), and DNA conformational energies were calculated as previously described (Olson et al., 1998). The minor groove width was measured directly as the distance between phosphorus atoms on opposite strands staggered by three base pairs.

Acknowledgments

We wish to thank Younchang Kim and the staff at SBC-CAT 19-ID (APS) and James Holton and the staff at ALS beamline 8.3.1 for excellent technical support during data collection. We would also like to thank Dave Stuart for helpful discussions and critical reading of this manuscript. This work was supported by a grant from the Canadian Institutes of Health Research (CIHR) to J.N.M.G., and funding from the Alberta Synchrotron Institute for synchrotron access. J.N.M.G acknowledges the support of a Canada Research Chair. J.S.L. was supported by a CIHR Canadian Graduate Scholarship and an Alberta Heritage Foundation for Medical Research Studentship.

Received: September 26, 2005
Revised: October 27, 2005
Accepted: November 13, 2005
Published online: March 14, 2006

References

Bravo, J., Li, Z., Speck, N.A., and Warren, A.J. (2001). The leukemia-associated AML1 (Runx1)-CBF beta complex functions as a DNA-induced molecular clamp. Nat. Struct. Biol. *8*, 371–378.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., and Herskowitz, I. (1998). The transcriptional program of sporulation in budding yeast. Science *282*, 699–705.

Chu, S., and Herskowitz, I. (1998). Gametogenesis in yeast is regulated by a transcriptional cascade dependent on Ndt80. Mol. Cell *1*, 685–696.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).

Fingerman, I.M., Sutphen, K., Montano, S.P., Georgiadis, M.M., and Vershon, A.K. (2004). Characterization of critical interactions between Ndt80 and MSE DNA defining a novel family of Ig-fold transcription factors. Nucleic Acids Res. *32*, 2947–2956.

Hepworth, S.R., Ebisuzaki, L.K., and Segall, J. (1995). A 15-base-pair element activates the SPS4 gene midway through sporulation in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *15*, 3934–3944.

Hepworth, S.R., Friesen, H., and Segall, J. (1998). NDT80 and the meiotic recombination checkpoint regulate expression of middle sporulation-specific genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *18*, 5750–5761.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A *47 (Pt. 2)*, 110–119.

Lamoureux, J.S., Maynes, J.T., and Glover, J.N. (2004). Recognition of 5′-YpG-3′ sequences by coupled stacking/hydrogen bonding interactions with amino acid residues. J. Mol. Biol. *335*, 399–408.

Lamoureux, J.S., Stuart, D., Tsang, R., Wu, C., and Glover, J.N. (2002). Structure of the sporulation-specific transcription factor Ndt80 bound to DNA. EMBO J. *21*, 5721–5732.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lu, X.J., Shakked, Z., and Olson, W.K. (2000). A-form conformational motifs in ligand-bound DNA structures. J. Mol. Biol. *300*, 819–840.

Mitchell, A.P. (1994). Control of meiotic gene expression in *Saccharomyces cerevisiae*. Microbiol. Rev. *58*, 56–70.

Montano, S.P., Cote, M.L., Fingerman, I., Pierce, M., Vershon, A.K., and Georgiadis, M.M. (2002). Crystal structure of the DNA-binding domain from Ndt80, a transcriptional activator required for meiosis in yeast. Proc. Natl. Acad. Sci. USA *99*, 14041–14046.

Nelson, H.C., Finch, J.T., Luisi, B.F., and Klug, A. (1987). The structure of an oligo(dA).oligo(dT) tract and its biological implications. Nature *330*, 221–226.

Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M., and Zhurkin, V.B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. Proc. Natl. Acad. Sci. USA *95*, 11163–11168.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In Methods in Enzymology, C.W. Carter, Jr. and R.M. Sweet, eds. (New York: Academic Press), pp. 307–326.

Ozsarac, N., Straffon, M.J., Dalton, H.E., and Dawes, I.W. (1997). Regulation of gene expression during meiosis in *Saccharomyces cerevisiae*: SPR3 is controlled by both ABFI and a new sporulation control element. Mol. Cell. Biol. *17*, 1152–1159.

Pak, J., and Segall, J. (2002). Regulation of the premiddle and middle phases of expression of the NDT80 gene during sporulation of *Saccharomyces cerevisiae*. Mol. Cell. Biol. *22*, 6417–6429.

Pierce, M., Benjamin, K.R., Montano, S.P., Georgiadis, M.M., Winter, E., and Vershon, A.K. (2003). Sum1 and Ndt80 proteins compete for binding to middle sporulation element sequences that control meiotic gene expression. Mol. Cell. Biol. *23*, 4814–4825.

Pierce, M., Wagner, M., Xie, J., Gailus-Durner, V., Six, J., Vershon, A.K., and Winter, E. (1998). Transcriptional regulation of the SMK1 mitogen-activated protein kinase gene during meiotic development in Saccharomyces cerevisiae. Mol. Cell. Biol. *18*, 5970–5980.

Prive, G.G., Heinemann, U., Chandrasegaran, S., Kan, L.S., Kopka, M.L., and Dickerson, R.E. (1987). Helix geometry, hydration, and G.A mismatch in a B-DNA decamer. Science *238*, 498–504.

Ptashne, M. (1986). A genetic switch: gene control and phage [lamda]. (Palo Alto, CA: Blackwell Scientific Publications & Cell Press).

Tahirov, T.H., Inoue-Bungo, T., Morii, H., Fujikawa, A., Sasaki, M., Kimura, K., Shiina, M., Sato, K., Kumasaka, T., Yamamoto, M., et al. (2001). Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta. Cell *104*, 755–767.

Tung, K.S., Hong, E.J., and Roeder, G.S. (2000). The pachytene checkpoint prevents accumulation and phosphorylation of the

meiosis-specific transcription factor Ndt80. Proc. Natl. Acad. Sci. USA *97*, 12187–12192.

Wang, W., Cherry, J.M., Nochomovitz, Y., Jolly, E., Botstein, D., and Li, H. (2005). Inference of combinatorial regulation in yeast transcriptional networks: a case study of sporulation. Proc. Natl. Acad. Sci. USA *102*, 1998–2003.

Xie, J., Pierce, M., Gailus-Durner, V., Wagner, M., Winter, E., and Vershon, A.K. (1999). Sum1 and Hst1 repress middle sporulation-specific gene expression during mitosis in *Saccharomyces cerevisiae*. EMBO J. *18*, 6448–6454.

Xu, L., Ajimura, M., Padmore, R., Klein, C., and Kleckner, N. (1995). NDT80, a meiosis-specific gene required for exit from pachytene in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *15*, 6572–6581.

Xu, L., Weiner, B.M., and Kleckner, N. (1997). Meiotic cells monitor the status of the interhomolog recombination complex. Genes Dev. *11*, 106–118.

Yoon, C., Prive, G.G., Goodsell, D.S., and Dickerson, R.E. (1988). Structure of an alternating-B DNA helix and its relationship to A-tract DNA. Proc. Natl. Acad. Sci. USA *85*, 6332–6336.

Accession Numbers

Coordinates have been deposited in the PDB (http://www.rcsb.org) with accession codes 2EUV (vG1A/A9T), 2ETW (vG1C), 2EUX (vA4G), 2EUW (mA4T), 2EUZ (mC5T), 2EVF (mA6T), 2EVG (mA7T), 2EVH (mA7G), 2EVI (mA8T), and 2EVJ (mA9C).