

Structure of the sporulation-specific transcription factor Ndt80 bound to DNA

Jason S.Lamoureux, David Stuart,
Roger Tsang, Cynthia Wu and
J.N.Mark Glover¹

Department of Biochemistry, University of Alberta, Edmonton, AB,
T6G 2H7, Canada

¹Corresponding author
e-mail: mark.glover@ualberta.ca

Progression through the middle phase of sporulation in *Saccharomyces cerevisiae* is promoted by the successful completion of recombination at the end of prophase I. Completion of meiotic recombination allows the activation of the sporulation-specific transcription factor Ndt80, which binds to a specific DNA sequence, the middle sporulation element (MSE), and activates ~150 genes to enable progression through meiosis. Here, we isolate the DNA-binding domain of Ndt80 and determine its crystal structure both free and in complex with an MSE-containing DNA. The structure reveals that Ndt80 is a member of the Ig-fold family of transcription factors. The structure of the DNA-bound form, refined at 1.4 Å, reveals an unexpected mode of recognition of 5'-pyrimidine-guanine-3' dinucleotide steps by arginine residues that simultaneously recognize the 3'-guanine base through hydrogen bond interactions and the 5'-pyrimidine through stacking/van der Waals interactions. Analysis of the DNA-binding affinities of MSE mutants demonstrates the central importance of these interactions, and of the AT-rich portion of the MSE. Functional similarities between Ndt80 and the *Caenorhabditis elegans* p53 homolog suggest an evolutionary link between Ndt80 and the p53 family.

Keywords: MSE/Ndt80/sporulation/transcription factor

Introduction

Meiosis in *Saccharomyces cerevisiae* is a tightly regulated process that couples the replication and segregation of chromosomes with morphological changes that accompany the development of the gamete (Roeder, 1997). This process is temporally regulated by at least four distinct waves of gene expression: early, middle, mid-late and late (Chu *et al.*, 1998).

In *S.cerevisiae*, the middle phase of meiosis initiates with the successful completion of recombination and segregation of paired, homologous chromosomes at the end of prophase I. This process is monitored to ensure that meiotic recombination is completed prior to synaptonemal complex dissolution and chromosome separation at anaphase of meiosis I (Xu *et al.*, 1997). A key transcription factor that regulates this process is Ndt80 (Xu *et al.*, 1995), which activates the transcription of >150 genes during the

mid-phase of sporulation (Chu and Herskowitz, 1998; Chu *et al.*, 1998; Hepworth *et al.*, 1998). Many of the genes activated by Ndt80 are required for the disassembly of the synaptonemal complex (e.g. Ubc9) or for further progression through the developmental process, such as several of the B-type cyclins (Chu and Herskowitz, 1998; Chu *et al.*, 1998). There is a subsequent wave of transcription (the 'mid-late' wave), which also appears to be dependent on Ndt80, but may also be regulated by repressor molecules that are responsible for the delay in the expression of these genes relative to the middle genes (Chu *et al.*, 1998).

Ndt80-dependent activation of middle gene expression is dependent on a 9 bp regulatory sequence within a few hundred base pairs of the start site of transcription (Hepworth *et al.*, 1995; Ozsarac *et al.*, 1997). A consensus sequence for this site, called the middle sporulation element (MSE), has been derived based on the comparison of the upstream regulatory regions of several Ndt80-activated genes (Chu and Herskowitz, 1998; Chu *et al.*, 1998). Electrophoretic mobility shift experiments using affinity-purified Ndt80 from yeast, or recombinant Ndt80 purified from *Escherichia coli*, have demonstrated that Ndt80 specifically binds this regulatory sequence (Chu and Herskowitz, 1998).

Interestingly, several isolated MSEs from middle genes act not only as positive regulatory sites during the mid phase of sporulation, but also repress gene expression during vegetative growth through the actions of two proteins, Sum1 and Hst1. Sum1 binds specifically to MSEs and the apparent affinity of Sum1 for a given MSE roughly correlates with the ability of the MSE to repress transcription. These findings led to the suggestion that the relative affinities of an MSE for either Sum1 or Ndt80 could govern the precise timing and level of activation of a gene under its control (Xie *et al.*, 1999; Lindgren *et al.*, 2000). Consistent with this model, the levels of Sum1 protein drop during the mid-phase of sporulation, at a time when Ndt80 levels rise and the middle sporulation genes are activated (Chu *et al.*, 1998; Lindgren *et al.*, 2000).

NDT80 was initially isolated in a genetic screen for mutants that are defective in progression through meiosis (Xu *et al.*, 1995). *ndt80* mutants were found to arrest during the pachytene stage of prophase I, suggesting that *NDT80* is involved in, or is regulated by, the successful completion of meiotic recombination during pachytene. The behavior of *ndt80 spo11* or *ndt80 rad50* double mutants suggested that *NDT80* is not required for recombination since neither the *spo11* nor *rad50* mutations, which cause a bypass of meiotic recombination, alleviated *ndt80*-dependent arrest of meiosis. In addition, a *dmc1* mutation, which blocks completion of meiotic recombination and thereby leads to activation of the recombination checkpoint and pachytene arrest (Lydall *et al.*, 1996), blocks *NDT80*-activated gene expression. Middle gene expression can be restored by a

rad17 mutation, which allows bypass of the recombination checkpoint (Chu and Herskowitz, 1998; Hepworth *et al.*, 1998). These results together suggest that *NDT80*, like *CDC28*, is required for progression from prophase to metaphase I and is regulated by the recombination checkpoint (Xu *et al.*, 1995; Chu and Herskowitz, 1998; Hepworth *et al.*, 1998).

It has been suggested that the recombination checkpoint-dependent regulation of *NDT80* is post-transcriptional (Chu and Herskowitz, 1998; Tung *et al.*, 2000). Support for this idea comes from the finding that Ndt80 is post-translationally modified during meiosis, as assayed by changes in the gel electrophoretic mobility of the protein throughout meiosis. Furthermore, in *dmc1* mutants that do not proceed through the meiotic recombination checkpoint, middle gene expression is completely blocked, in spite of the fact that *NDT80* is still transcribed (Chu and Herskowitz, 1998; Hepworth *et al.*, 1998).

To understand how Ndt80 binds MSE DNA, we have used genetic and biochemical approaches to delineate the DNA-binding domain of Ndt80 and have determined its structure bound to an MSE-containing oligonucleotide at 1.4 Å resolution. The structure reveals that Ndt80 is a member of the Ig-fold family of transcription factors, and recognizes DNA in a manner similar to other members of the family such as p53, NF-κB and NFAT. The integrity of the DNA-binding domain is shown to be essential for transcriptional activation and normal progression through sporulation. The structure reveals that two conserved 5'-YpG-3' dinucleotide steps are recognized by arginine residues, which simultaneously hydrogen bond to the guanine base and stack with the 5'-pyrimidine base, while the poly(A)-poly(T) portion of the MSE is recognized by minor-groove interactions. A detailed analysis of the effects of mutations within the MSE on Ndt80-binding affinity reveals that both the 5'-YpG-3' dinucleotide steps and the poly(A)-poly(T) tract are critical specificity determinants for Ndt80 binding.

Results

Elucidation of the Ndt80 DNA-binding domain

Ndt80 bears little detectable sequence similarity to other proteins that might shed light on its structure or function. Using deletion analysis, we isolated an N-terminal DNA-binding fragment of Ndt80 [Ndt80(1-349)] and, using limited proteolysis, we were able to show that residues 59-340 adopt a proteolytically resistant conformation in the absence of DNA, while the N-terminal residues 1-58 only become protected against proteolysis when bound to DNA (Materials and methods). We next compared the abilities of full-length Ndt80, Ndt80(1-340) and Ndt80(59-340) to specifically bind to an MSE-containing DNA in an electrophoretic mobility shift assay (EMSA). The results of this experiment (Figure 1A) show that while all three proteins are capable of binding the MSE, only full-length Ndt80 and Ndt80(1-340) form a complex that is resistant to challenge with a non-specific competitor DNA [poly(dI-dC)]. Thus, residues 1-58 of Ndt80 are essential for sequence-specific recognition of the MSE and we conclude that Ndt80(1-340) contains a minimal DNA-binding domain that contains all the protein elements necessary for specific MSE recognition.

To confirm that this region of the protein forms an active MSE DNA-binding domain *in vivo*, we generated a series of N-terminal truncation mutants of Ndt80 and placed them under the regulation of the endogenous *NDT80* promoter. These mutant constructs were used to replace the endogenous *NDT80* in an otherwise wild-type *a/α* diploid yeast strain. The functions of the mutant versions of Ndt80 were tested by inducing the cells to sporulate and then assaying the expression of an MSE-regulated *lacZ* reporter gene. Cells carrying wild-type and mutant versions of *NDT80* were induced to initiate sporulation, and β-galactosidase activity was monitored 6 h after the induction of sporulation. This corresponds to the time frame in which most Ndt80-regulated genes are activated (Chu *et al.*, 1998). Full-length Ndt80 was most effective in promoting MSE-regulated reporter gene activity in this assay (Figure 1B). A 20 amino acid N-terminal truncation had only a modest effect on the ability of Ndt80 to activate an MSE-regulated reporter gene, while larger N-terminal deletions had an increasingly deleterious effect on the ability of Ndt80 to activate the reporter gene. Truncations of >60 amino acids profoundly reduced the ability of Ndt80 to activate the MSE-regulated reporter gene. These data display a direct correlation with the ability of the Ndt80 N-terminal deletion mutants to support the completion of sporulation. Truncations of ≥60 amino acids led to a significant reduction in spore formation. This observation is consistent with the idea that many of the middle sporulation genes regulated by Ndt80 are not effectively activated by the Δ60 mutant. Thus, the N-terminal 60 residues of Ndt80 are essential for sequence-specific MSE recognition, MSE-dependent transcriptional activation and normal progression through sporulation.

Binding affinity of Ndt80 for mutant MSEs

Ndt80 activates >150 distinct genes to allow the progression through meiotic prophase I (Chu *et al.*, 1998). A comparison of the upstream regulatory sequences of a number of genes that are activated by Ndt80 revealed a 9 bp MSE consensus sequence: 5'-gNCRCAA(A/T)-3' (where g represents a non-conserved guanine nucleotide, R a purine nucleotide and N a non-conserved position). Here we label bases 1-9 in this strand of the MSE in the 5'-3' direction; the complementary bases in the opposite strand are given the same numbers but are distinguished with a prime (Figures 4A and 6). To understand the relative importance of individual base pairs within the MSE to Ndt80-binding affinity, we created a set of MSE-containing DNA oligonucleotides in which each position within the MSE was individually replaced. The DNAs were based on the MSE from the *SPS4* promoter, which exactly matches the MSE consensus, and has previously been shown to bind Ndt80 and activate transcription (Hepworth *et al.*, 1995; Chu and Herskowitz, 1998). Ndt80 binding to the mutated DNAs was assayed by EMSA experiments (see Materials and methods). A graph summarizing the relative effects of these mutations on Ndt80-binding affinity is shown in Figure 2. The results show that while the substitution of any of the consensus MSE base pairs leads to a reduction in Ndt80-binding affinity, by far the most dramatic substitution was the replacement of the C5-G5' pair with G5-C5', which decreases the affinity of binding >100-fold. Substitution of

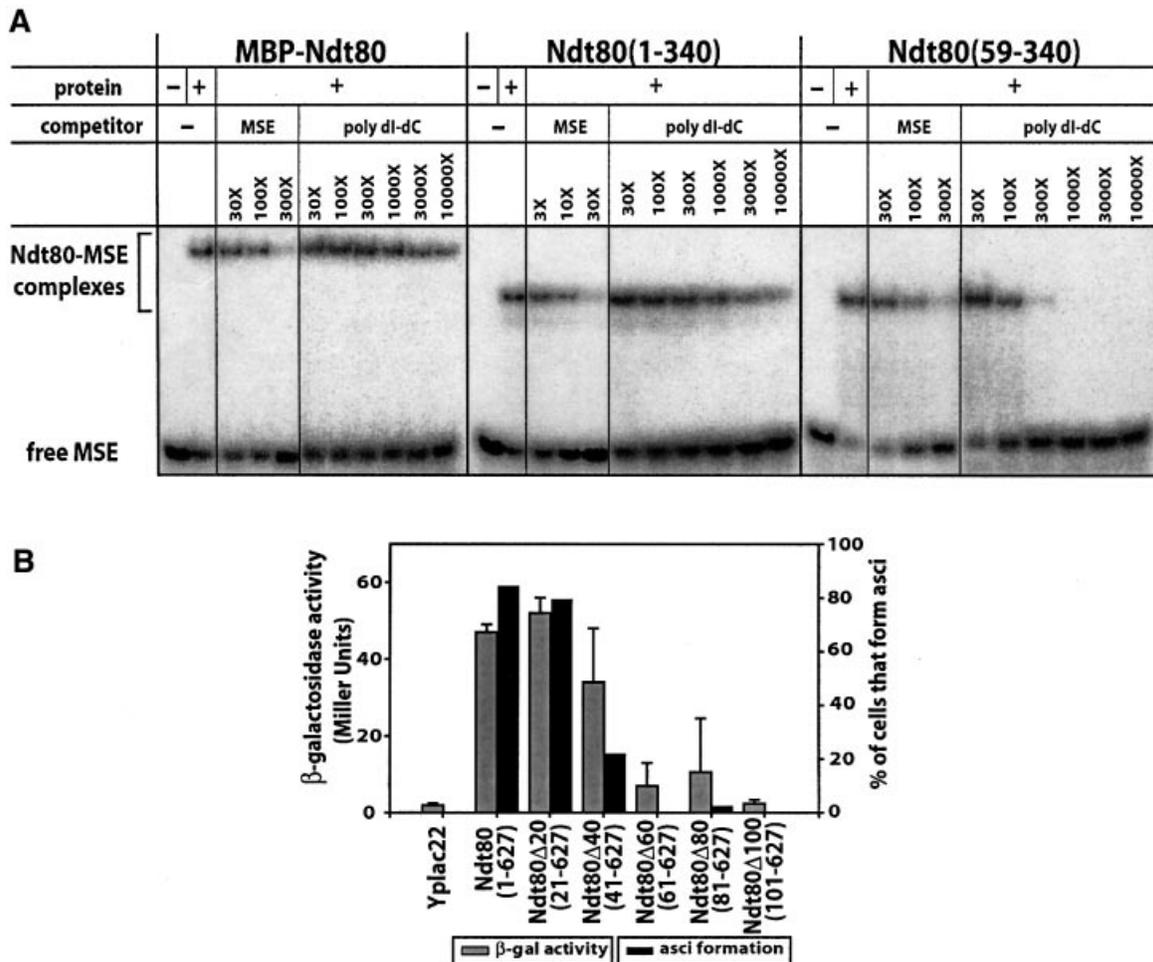


Fig. 1. Definition of the Ndt80 DNA-binding domain. (A) Purified recombinant MBP-Ndt80, Ndt80(1–340) or Ndt80(59–340) was bound to 32 P-labeled MSE DNA and challenged with either the unlabeled MSE DNA as a specific competitor or poly(dI-dC) as non-specific competitor at a variety of molar ratios of competitor to 32 P-labeled MSE DNA, as indicated. Note that an \sim 10-fold lower concentration of Ndt80(1–340) than MBP-Ndt80 was used to bind 32 P-labeled MSE DNA. As a result, \sim 10-fold less of either cold DNA is required to compete the Ndt80(1–340)–MSE complex, compared with the MBP-Ndt80–MSE complex. (B) Summary of the *in vivo* N-terminal truncation mutants. A series of mutants encoding the indicated Ndt80 fragments were transformed into yeast cells lacking endogenous Ndt80. Six hours after shift to sporulation medium, MSE-dependent transcription was assayed using a *lacZ* gene under the control of the *SPS4* upstream sequence containing an MSE. The gray bars represent β -galactosidase activity in Miller units. After 24 h, sporulation was monitored by microscopic examination of the culture to score for ascus formation (black bars).

the C3–G3' base pair showed the next most significant effect, followed by the substitution of the A–T pairs at positions 4 and 6, immediately 3' to the conserved C–G base pairs. A–T substitutions at positions 6–8 also resulted in significant decreases in binding affinity, suggesting that Ndt80 specifically recognizes the poly(A)–poly(T) tract at the 3' end of the MSE, and not merely an AT-rich region. To further probe the determinants of specificity within the poly(A)–poly(T) tract, we individually substituted each of the thymine bases within this region with uracils, effectively replacing the 5-methyl group with a hydrogen atom in the modified base. Of these substitutions, only replacement of T6' with U caused a significant decrease in Ndt80-binding affinity, indicating a significant role for the 5-methyl group at this position.

Overall structure

We have crystallized and determined the structure of Ndt80(1–340) bound to a 14 bp DNA duplex containing a

consensus MSE sequence derived from the *SPS4* gene, as well as the structure of Ndt80(59–340) in the absence of DNA (Materials and methods; Table I; Figure 3). The proteolytically resistant core domain reveals a central β -sandwich structure characteristic of an s-type Ig fold (Bork *et al.*, 1994) (Figure 4B). An alignment of the free and bound forms of Ndt80 reveals little structural change upon DNA binding (r.m.s.d. = 0.8 Å for all C_{α} atoms in common), other than the ordering of several loops that directly contact DNA (see below). The β -sandwich contains a three-stranded sheet composed of strands a, b and e, packed against a four-stranded sheet composed of strands c', c, f and g, labeled according to the standard nomenclature (Bork *et al.*, 1994). Each sheet of the β -sandwich contains an additional β -strand, as well as a variety of peripheral secondary structure elements, which are numbered according to their occurrence in the primary sequence (Figure 4B and C). A search of the protein structure database using either the DALI (Holm and

Sander, 1997) or VAST programs (Madej *et al.*, 1995) revealed that many of the proteins that are most similar to Ndt80 at the structural level are transcription factors that bind DNA through domains that also contain an s-type Ig fold: members of the p53 (Cho *et al.*, 1994), Rel/NF- κ B (Ghosh *et al.*, 1995; Muller *et al.*, 1995; F.E.Chen *et al.*, 1998; L.Chen *et al.*, 1998), STAT (Becker *et al.*, 1998; X.Chen *et al.*, 1998), CBF/Runx (Bravo *et al.*, 2001; Tahirov *et al.*, 2001) and T-box families (Muller and Herrmann, 1997) (for a recent review of the Ig-fold family of transcription factors, see Rudolph and Gergen, 2001). Ndt80 is the first non-metazoan member of the Ig-fold family of transcription factors.

The structure of the complex of Ndt80 bound to the MSE has been refined to 1.4 Å resolution, revealing a large and complex protein–DNA interface (Figures 4A, 5 and 6). Ndt80 interacts with both the major and minor grooves of the DNA through six distinct regions extending from one end of the β -sandwich. The face of the Ig fold that contacts DNA is the same as seen in other Ig-fold transcription factors and three of the DNA-contact regions are conserved. The DNA adopts typical B-DNA geometry and has no dramatic bends or kinks. Like the DNA sequences contacted by many of the other Ig-fold transcription factors, the MSE contains a 5' GC-rich region that is specifically recognized through interactions with amino acid side chains via the DNA major groove, and a 3' AT-rich region that is largely contacted via the DNA minor groove.

Comparison with other Ig-fold transcription factors

Not only does Ndt80 share a core β -sandwich topology with the other Ig-fold transcription factors, but it also binds DNA using many of the same loops employed by the other members of the family (Figures 3 and 4). For example, the a–b loop in Ndt80 contains a β -hairpin that is oriented perpendicular to the plane of the β -sandwich to interact both with the DNA major groove and with the N- and C-terminal extensions that also contact DNA. The a–b loop in p53 (Cho *et al.*, 1994), as well as in the NF- κ B p50-like proteins (Ghosh *et al.*, 1995; Muller *et al.*, 1995) and NFAT (L.Chen *et al.*, 1998), also adopts β -hairpin conformations, although in the case of the NF- κ B and NFAT proteins, the hairpin points in the opposite direction to that found in Ndt80, so that more extensive contacts with the DNA can be made. The e–f loop in Ndt80 makes contacts to the backbone of the DNA and, in other Ig-fold transcription factors, this loop plays a similar role, and also interacts directly with the base pairs via the minor groove. The e–f loop is quite variable in size and structure, containing a large helical insert characteristic of the NF- κ B p50 family. The C-terminal tail of Ndt80 extends away from the body of the protein to make critical contacts with the DNA major groove. The C-terminal extension of several other Ig-fold transcription factors also interacts with the DNA major groove, although the conformation of this loop is quite variable. For example, in p53, the extension adopts a helical conformation that packs against the a–b loop and interacts with the DNA major groove, while in NF- κ B p50 and AML1/Runx1, the C-terminal tail extends away from the Ig fold to contact DNA.

Ndt80 also contains several large loop regions that are unique to this protein. Ndt80 contains an N-terminal

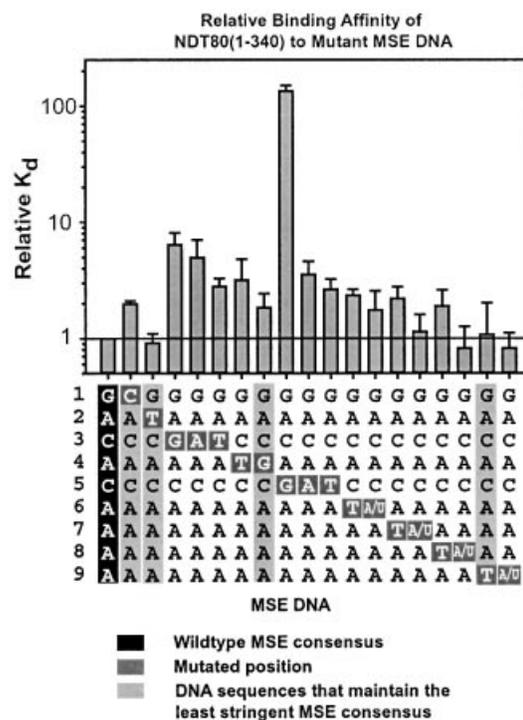


Fig. 2. Mutation of the MSE reduces Ndt80-binding affinity. The dissociation constants (K_d) for the binding of Ndt80(1–340) to wild-type and mutated MSE DNAs were determined by EMSA. Average K_d values relative to the consensus *SPS4* MSE are plotted, together with SDs derived from at least three independent experiments for each mutant.

extension (residues 1–89), which consists of a β -hairpin and a loop that contacts the DNA minor groove and is essential for sequence-specific recognition of the MSE. Residues 1–32 are not visible in the electron density, and these residues are not required for sequence-specific recognition of the MSE (data not shown), or for Ndt80-dependent transcriptional activation or progression through meiosis (Figure 1). The finding that the loop region (residues 50–60) is sensitive to proteolysis in the unbound state, but not in complex with DNA, suggests that this region is conformationally flexible and adopts its ordered three-dimensional structure only upon binding the MSE. Residues 70–80 constitute a β -strand that is unique to Ndt80 and binds to the g strand of the c'–c–f–g β -sheet.

The loops linking the b–c, c–c', c'–e and f–g strands of Ndt80 tend to be larger and more complex than equivalent loops found in the other Ig-fold transcription factors. The large c–c' loop contains a β -hairpin structure and makes novel, sequence-specific contacts to the DNA major groove. The c'–e loop contains a helix–loop–helix insert that recognizes the narrow minor groove of the AT-rich portion of the MSE. The f–g loop contains an extra strand (β 13) that extends the c'–c–f–g sheet.

Overall, the Ndt80 DNA-binding domain makes more extensive contacts with the MSE than is observed between other, single Ig domains and their target DNAs. In all, 2600 Å² of solvent-accessible surface area are buried in the Ndt80–DNA interface. This is roughly twice the area buried by other Ig-fold transcription factors. This might explain why many of the other Ig-fold transcription factors only bind DNA with high affinity as dimers (as is the case

Table I. Summary of X-ray experiments

Crystal data	Ndt80(1–340) + MSE DNA			Ndt80(59–340)			
<i>a</i> (Å)	70.13			35.88			
<i>b</i> (Å)	78.81			41.93			
<i>c</i> (Å)	161.39			163.91			
Space group	C222 ₁			P2 ₁ 2 ₁ 2 ₁			
Data collection	Native	MAD λ ₁	MAD λ ₂	MAD λ ₃	MAD λ ₁	MAD λ ₂	MAD λ ₃
Wavelength (Å)	0.98008	0.97936	0.97922	0.93218	0.9793	0.9789	0.9563
Resolution	100–1.4	100–1.7	100–1.7	100–1.7	40–2.2	40–2.2	40–2.2
Reflections							
Total	3 338 673	218 954	219 186	220 867	141 688	141 567	144 249
Unique	88 384	49 236	49 277	49 065	12 360	12 286	12 292
Completeness	99.8 (97.8)	98.6 (87.3)	98.4 (82.8)	98.1 (90.1)	99.1 (100)	97.8 (100)	98.1 (100)
<i>I</i> / <i>σ</i>	35 (3.0)	18.5 (3.5)	18.1 (3.1)	18.1 (3.9)	19.9 (7.3)	17.5 (5.4)	20.9 (7.9)
<i>R</i> _{sym}	5.1 (49.3)	4.5 (23.8)	5.3 (29.7)	4.5 (30.4)	8.6 (25.4)	9.9 (32.7)	7.6 (22.1)
Redundancy	38 (5.2)	4.5 (2.1)	4.5 (2.1)	4.5 (2.6)	11.5 (5.7)	10.5 (5.2)	11.7 (5.4)
Refinement	Ndt80(1–340) + MSE DNA			Ndt80(59–340)			
<i>R</i> _{cryst} / <i>R</i> _{free}	19.4/20.6			22.6/26.8			
R.m.s.d. bonds (Å)	0.011			0.006			
R.m.s.d. angles (°)	1.52			1.28			
Ramachandran plot (% residues in region)							
Favored	90.6			85.3			
Allowed	8.2			12.4			
Generously allowed	0.8			1.8			
Disallowed	0.4			0.5			

Values in parentheses are statistics for the highest resolution shell [1.42–1.40 Å for native and 1.73–1.70 Å for MAD data of Ndt80(1–340), and 2.25–2.20 Å for all data sets of Ndt80(59–340)].

$R_{\text{sym}} = 100 \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, for the intensity of *i* observations of reflection *h*.

$R_{\text{cryst}} = \sum_h |F_o(h) - F_c(h)| / \sum_h |F_o(h)|$, where $F_o(h)$ and $F_c(h)$ are observed and calculated structure factors.

*R*_{free} is calculated with 5% of all reflections excluded from refinement stages using high-resolution data.

Note that Ndt80(59–340) was refined against the MAD λ₃ (high energy remote) data set as no high-quality native data set was collected.

of the Rel/NF-κB family), or in complex with other accessory proteins (as is the case for NFAT and AML1/Runx1 proteins) (Rudolph and Gergen, 2001).

Ndt80–MSE specificity: 5′-YpG-3′ recognition

All Ig-fold transcription factors make sequence-specific contacts to their target DNAs through arginine side chains that recognize the major-groove face of guanine bases through a pair of hydrogen bonds between the guanidinium group of the arginine and the N7 and O6 atoms of the guanine (Rudolph and Gergen, 2001). In Ndt80, the conserved C–G pairs at positions 3 and 5 are recognized in this way, as is the semi-conserved G–C pair at position 1 (Figures 5A, 6 and 7A). G1 is recognized by R326 from the C-terminal tail, while G3′ is recognized by R111 from the a–b loop and G5′ is recognized by R177 from the c–c′ loop. These contacts explain the fact that mutation of any of the conserved G–C base pairs within the MSE leads to a significant loss of Ndt80-binding affinity (Figure 2).

In addition to direct recognition of the G–C base pair, pyrimidine bases (thymine favored over cytosine) are conserved at the positions immediately 5′ to the guanine nucleotides at positions 3 and 5. Our structure reveals that Ndt80 recognizes the 5′-pyrimidine residues by taking advantage of the inherent flexibility of 5′-pyrimidine–purine-3′ dinucleotide steps (Olson *et al.*, 1998). This inherent flexibility is probably due to poor overlap

between adjacent base pairs in these steps. Comparison of the 5′-T4′pG3′-3′ and 5′-T6′pG5′-3′ steps from the Ndt80–MSE structure with a 5′-TpG-3′ step from the crystal structure of a free DNA of nearly identical sequence (Nelson *et al.*, 1987) reveals that the thymine bases in the Ndt80-bound structure are displaced by ~1.6 Å into the major groove, almost completely sacrificing stacking with the 3′ guanine (Figure 7). Instead, the thymines (especially their 5-methyl groups) make van der Waals contacts to the arginine hydrogen bonded to the guanine, as well as to the main chain of the arginine and the adjacent residue. This observation explains the fact that mutation of either of the thymines to adenine results in a significant loss of binding affinity to Ndt80 (Figure 2). The importance of van der Waals contact between the 5-methyl group of T6′ and R177 is demonstrated by the fact that substitution of T6′ with U results in a significant loss of Ndt80 binding, whereas T–U substitutions in the other positions of the poly(A)–poly(T) tract do not affect binding. In addition, T6′ is also contacted through the minor groove by P57, which also makes significant van der Waals contact with the otherwise exposed face of G5′. The fact that the T6′–G5′ step is recognized in a concerted manner via both the major and minor grooves probably explains the dramatic, 100-fold reduction in binding affinity when the C5–G5′ base pair is mutated to G–C (Figure 2).

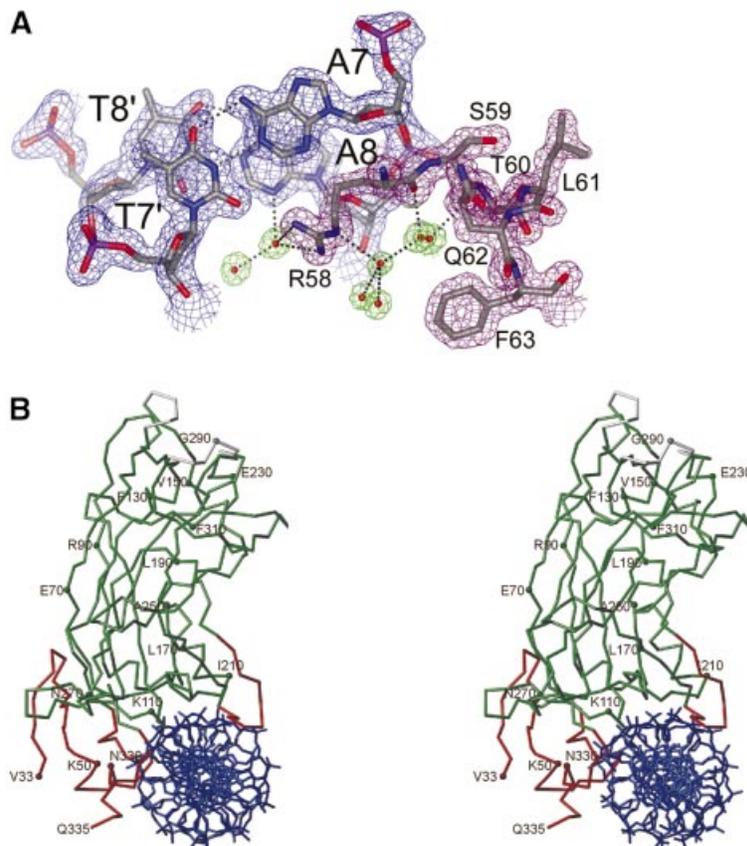


Fig. 3. Model of the Ndt80(1–340)–MSE complex at 1.4 Å resolution. (A) $2F_o - F_c$ electron density covering the DNA at the seventh and eighth position of the MSE is colored blue (1.5σ), density corresponding to protein is colored purple (1.5σ), while solvent density is colored green (1.4σ). This diagram shows the importance of the water molecules in the DNA–protein interface as well as the exclusion of G–C or C–G base pairs within the poly(A)–poly(T) tract due to steric clash of a 2-amino group of guanine with the R58 side chain. (B) C_α trace of Ndt80(1–340) bound to DNA (stereo view). The green trace indicates regions of the protein that could be modeled in both Ndt80(1–340) and Ndt80(59–340) DNA structures, while red regions are those that could not be modeled in the Ndt80(59–340) structure but are well ordered in the Ndt80(1–340) structure. The gray trace represents chain breaks in the Ndt80(1–340) complex structure.

Recently, *ab initio* energetic calculations and analysis of protein–DNA structures were used to suggest that arginines might interact most favorably with 5′-purine–guanine-3′ dinucleotide steps, where the 3′ guanine is recognized by hydrogen bonding, while the 5′-purine interacts with the arginine through cation– π interactions (Wintjens *et al.*, 2000; Rooman *et al.*, 2002). In Ndt80, however, a 5′-pyrimidine is clearly favored over a 5′-purine, probably because the conformational flexibility afforded by the 5′-pyrimidine–guanine-3′ allows a greater degree of pyrimidine–arginine contact.

A survey of the structures of other Ig-fold transcription factors bound to their cognate DNAs reveals that the AML1 (Runt domain) also employs coupled arginine–guanine hydrogen bonding and pyrimidine stacking to recognize 5′-pyrimidine–guanine-3′ steps in its consensus DNA (5′-TGTGGTT-3′) (Bravo *et al.*, 2001; Tahirov *et al.*, 2001). Interestingly, the specific recognition of 5-MeCpG dinucleotide steps over the unmethylated forms by MBD transcriptional repressors is mediated by a pair of arginine residues (Free *et al.*, 2001; Ohki *et al.*, 2001). Whether or not this recognition is also mediated by the coupled hydrogen bonding and stacking interactions between the arginines and the 5-MeCpG dinucleotide awaits the determination of such a structure at high resolution.

Ndt80–MSE specificity: minor-groove recognition

The poly(A)–poly(T) portion of the MSE (bp 6–9) is largely recognized by Ndt80 through interactions with the DNA minor groove (Figures 5B and 6). The N-terminal loop linking $\beta 2$ and $\beta 3$ enters the minor groove, and P57 makes van der Waals contacts with the minor-groove face of the A6–T6′ base pair (Figures 5B, 6 and 7A). The next residue in this loop, R58, makes van der Waals contacts to the A7–T7′ pair in the minor groove, and the R58 guanidinium group hydrogen bonds with T8′, as well as a water that hydrogen bonds to the A9–T9′ pair (Figures 3A, 5B and 6). Substitution of any of these A–T pairs with a G–C pair would result in a steric clash between the 2-amino group of the guanine base and the side chain. As a result of these interactions, the DNA minor groove near this end of the poly(A)–poly(T) region is significantly wider than that normally seen in B-DNA (~ 13 Å inter-strand phosphate distance, compared with an ~ 11 Å distance in B-DNA). However, toward the 3′ end of the poly(A)–poly(T) tract, the minor groove narrows significantly to ~ 9.5 Å, similar to the minor-groove width observed in the crystal structures of poly(A)–poly(T) in the absence of protein (Nelson *et al.*, 1987). The narrow minor groove in the vicinity of base pairs A7–T7′ to C11–G12′ is recognized by residues extending from a helix–loop–helix motif present in the c′–e

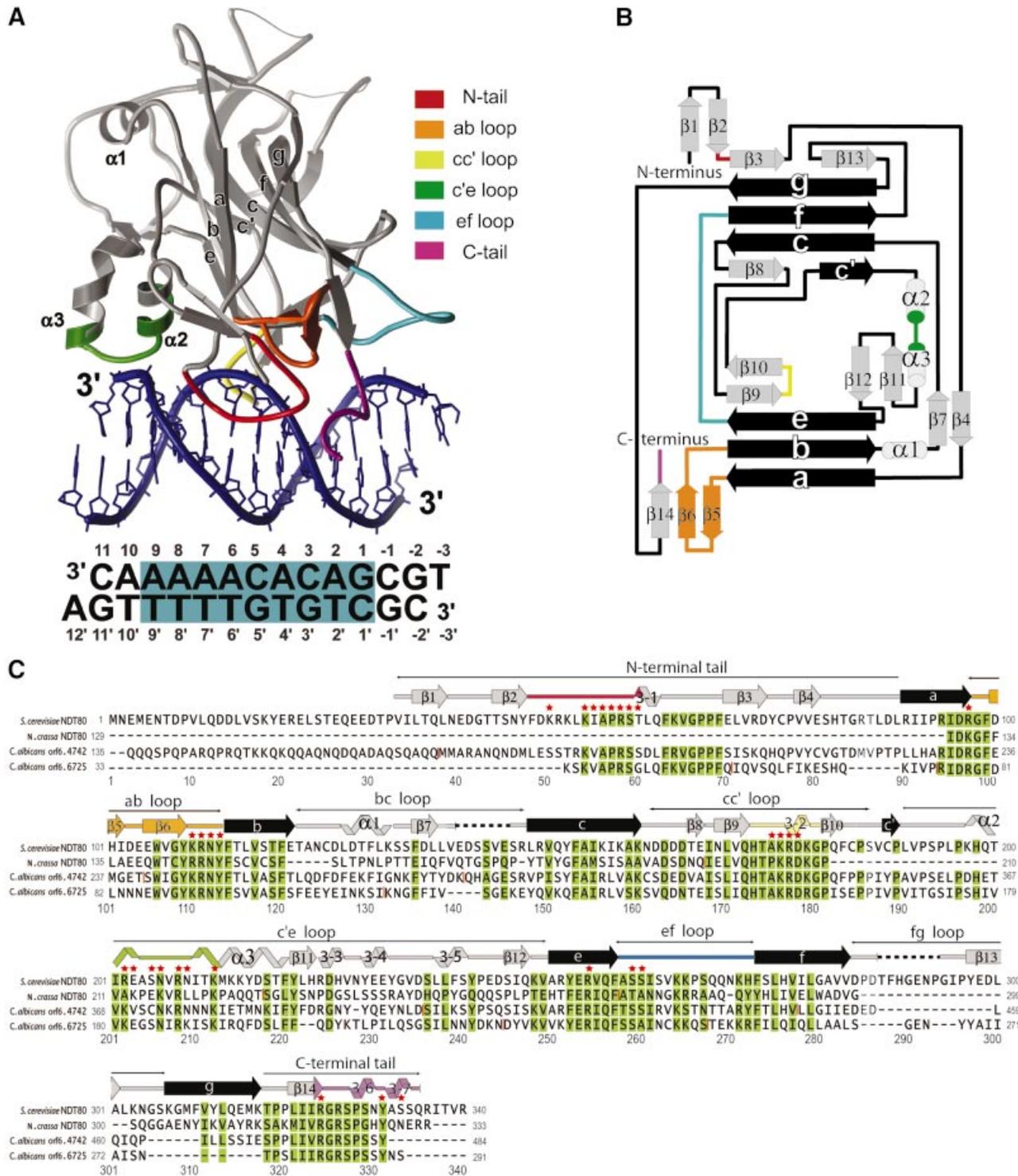


Fig. 4. Structure of Ndt80(1–340) bound to DNA. (A) Overview of the Ndt80(1–340)–MSE complex. The regions that contact DNA are rainbow-colored from N- to C-terminus, a color scheme that is maintained throughout the paper. Only the core Ig-fold β -strands and α -helices are labeled. Chain breaks are indicated by the thin coil. The DNA sequence is shown aligned with the structure and the MSE is highlighted. (B) Topology diagram of Ndt80. The conserved s-type Ig-fold strands are indicated in black, while the peripheral strands, in gray, are named according to their order in the primary sequence. (C) Secondary structure and sequence alignment of the DNA-binding domain of Ndt80. The secondary structure includes regions that form 3_{10} helices. *Saccharomyces cerevisiae* Ndt80 is aligned with sequences from other fungi that share high sequence similarity. The regions highlighted in green share $\geq 75\%$ sequence identity. Residues with stars above them are involved in DNA contacts. The red lines within the alignment indicate amino acid insertions that were omitted in order to maintain an ungapped sequence for *S. cerevisiae* Ndt80. The numbering along the bottom refers to the Ndt80 numbering, while the numbers on either side are for the indicated reading frame.

loop (Figure 5B). These residues make a number of direct and solvent-mediated hydrogen bond and electrostatic contacts to both strands across the narrow minor groove.

While our structure explains why A–T pairs are favored over G–C pairs in the 3' region of the MSE, it is less clear why substitution of any A–T pair with a T–A at positions

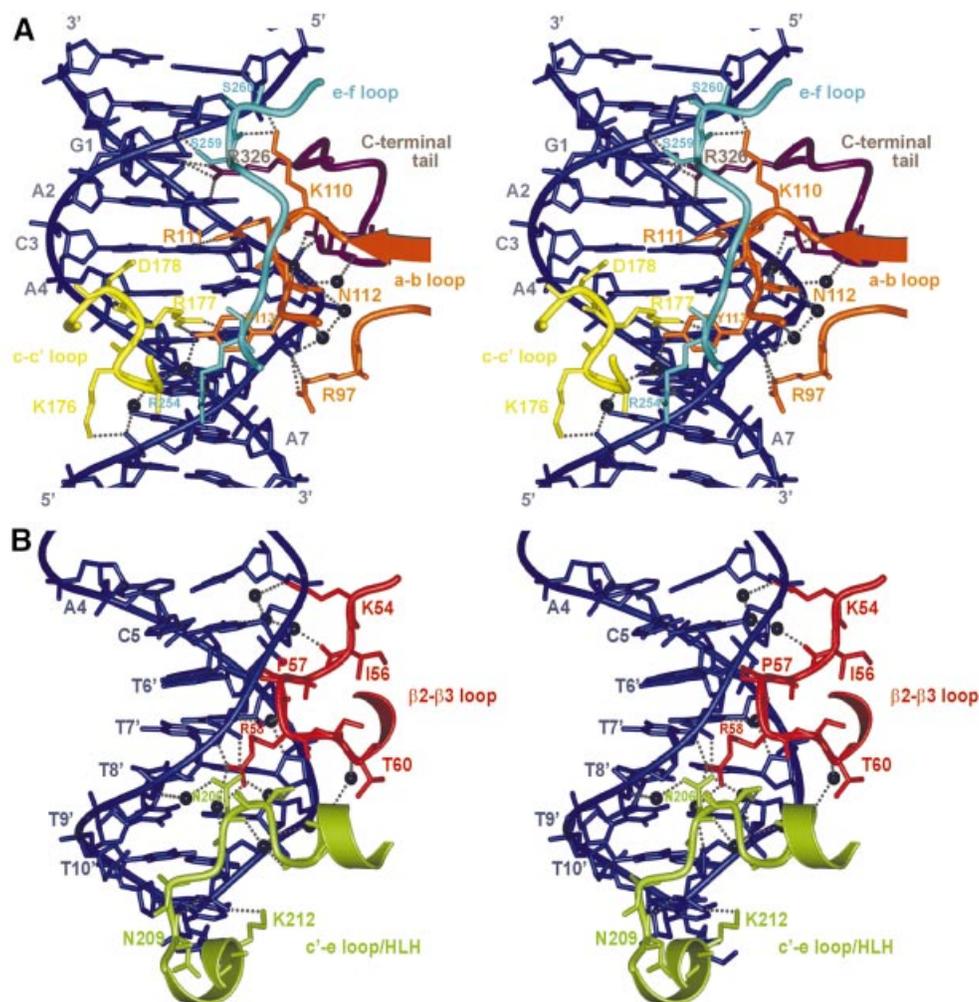


Fig. 5. Ndt80–DNA interface. (A) Stereo image of the major-groove interactions between Ndt80 and DNA. Involved in this interface are the a–b, c–c' and e–f loops, and the C-terminal tail. (B) Stereo image of the minor-groove interactions between Ndt80 and DNA. This interface involves the N-terminal tail (or β2–β3 loop) and the c'–e loop (HLH).

7–9 is unfavorable (Figure 2). Such substitutions would have little effect on either the disposition of hydrogen bonding groups or the width of the minor groove. A possible explanation may lie in the fact that poly(A)–poly(T) tracts are more rigid than A–T DNAs containing 5'-TpA-3' dinucleotide steps (Suzuki *et al.*, 1996). Because of this, an additional entropic cost might be associated with Ndt80 binding to a flexible, mixed A–T sequence over a more rigid poly(A)–poly(T) tract that is pre-set in a conformation appropriate for Ndt80 binding. The crystal structure of a free DNA with a sequence that is almost identical to the MSE DNA provides an approximation of the conformation of the MSE DNA in its free state (Nelson *et al.*, 1987). While the structure of this DNA is very similar to the conformation of the MSE DNA bound to Ndt80, interesting differences are present. In addition to the differences in minor-groove width mentioned above, the poly(A)–poly(T) tract in its unbound form also exhibits pronounced propeller twisting that at once enhances intra-strand base stacking and hydrogen bonding between adenine and thymine bases in adjacent base pairs. This propeller twisting is not observed in the Ndt80-bound form of the MSE, and suggests that protein binding may alter

the conformation of the DNA in subtle but significant ways that ultimately modulate the binding affinity.

Ndt80–MSE specificity: backbone contacts

In addition to the contacts described above, which play a key role in the sequence-specific recognition of the MSE, Ndt80 displays a complex network of electrostatic and hydrogen bonding interactions that anchor the protein to the DNA (Figures 5 and 6). The N- and C-terminal extensions, the a–b, c–c' and c'–e loops, which all mediate sequence-specific interactions with the DNA, as well as the e–f loop, all make contacts to the DNA backbone, many of which are mediated by well-ordered solvent molecules.

Ndt80 flexibility and DNA binding

A comparison of the DNA-bound form of Ndt80(1–340) with the unbound form of Ndt80(59–340) indicates that several of the DNA-contact regions only become structured upon interactions with DNA (Figure 3B). We observed no interpretable electron density for either the c–c' loop (corresponding to residues 175–184), or the C-terminal tail (residues 326–334) in the unbound form of

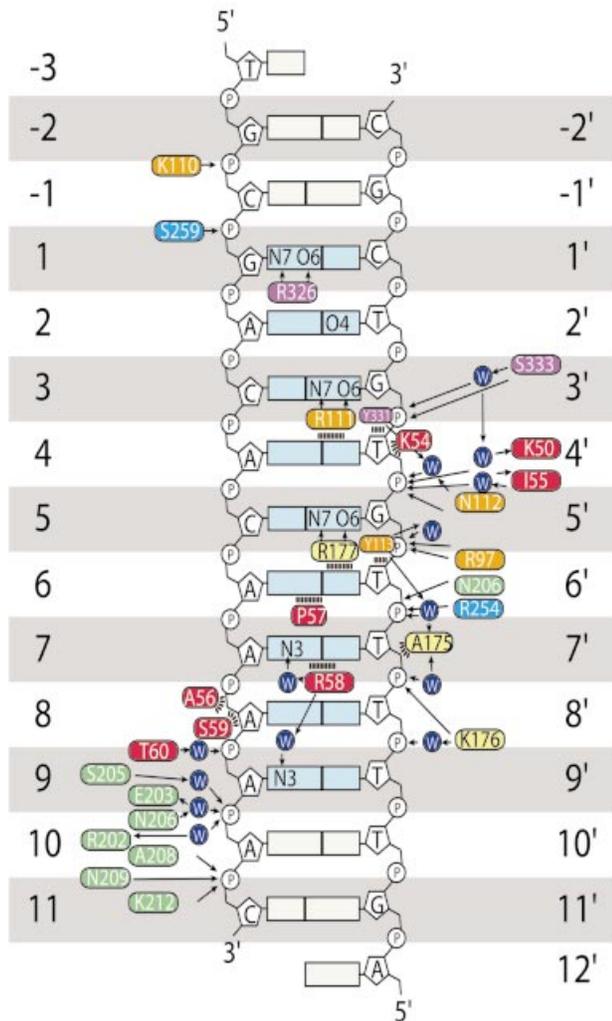


Fig. 6. Schematic of Ndt80–DNA interactions. The MSE bases are highlighted in blue and are numbered 1–9. The complementary strand is distinguished with a prime following the number. Protein residues are colored as described in Figure 4. Electrostatic and polar interactions are indicated with arrows, while van der Waals contacts are represented by hash marks. The blue circles are water molecules that are highly ordered in the DNA–protein interface.

Ndt80, indicating that these critical DNA-contact regions are flexible in the absence of DNA. However, the a–b, c–c', helix–loop–helix and e–f loops are ordered and adopt conformations that are essentially identical to those found in the DNA-bound form. In particular, R111 from the a–b loop that recognizes the 5'-TpG-3' dinucleotide step at position 3'-4' adopts a side-chain geometry almost identical to that seen in the DNA-bound form. Thus, these loops appear to be pre-aligned for interactions with the DNA, while the N- and C-terminal extensions, as well as the c–c' loop, are more flexible and probably only adopt stable structures when complexed with DNA.

Discussion

Interactions between Ndt80 and other transcription factors

While Ndt80 is the key transcription factor that ultimately allows the continuation of meiosis after the successful

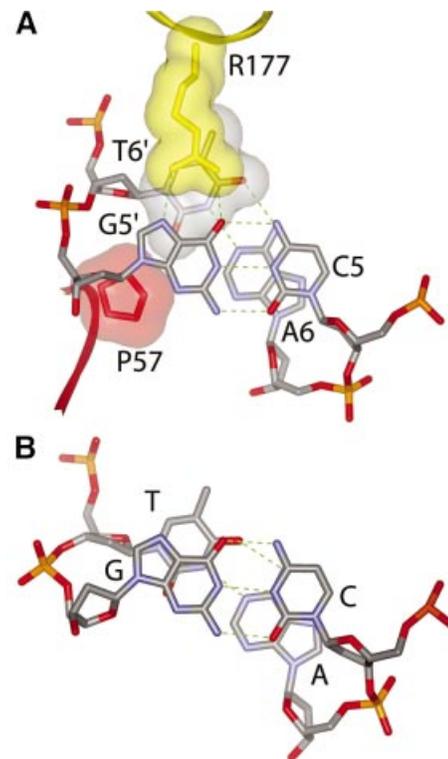


Fig. 7. 5'-YpG-3' recognition by Ndt80. (A) View down the DNA helix axis showing base stacking between the C–G base pair at position 5 and the A–T pair at position 6. van der Waals surface representations of R177 (yellow), P57 (red) and T6' (gray) are displayed and hydrogen bonding interactions are indicated in green. (B) An equivalent CG–AT is shown from the structure of a similar DNA sequence determined in the absence of bound protein (Nelson *et al.*, 1987). The view is such that the C–G base pairs in both panels are in identical orientations.

completion of recombination, other factors must also induce middle gene expression. Evidence for this comes from the finding that many middle genes are transcriptionally activated in Ndt80-deficient cells during meiosis, albeit to a lower level than in isogenic strains containing wild-type Ndt80 (Chu *et al.*, 1998). Part of this Ndt80-independent induction may be due to the release of repression imposed on middle genes by Sum1–Hst1 or other repressor complexes during vegetative growth and the early stages of meiosis (Xie *et al.*, 1999; Pijnappel *et al.*, 2001). Sum1 has been shown to interact with MSEs in a sequence-specific manner, probably utilizing its two AT-hook peptide motifs (Aravind and Landsman, 1998), which probably make sequence-specific contacts with the AT-rich portion of the MSE. Because Ndt80 makes extensive contacts with both the backbone and floor of the minor groove of the AT-rich region, it is very likely that the Ndt80 activator and the Sum1 repressor compete for the same DNA regulatory elements. This competition may ultimately help to regulate the precise timing and level of expression of key middle genes (Xie *et al.*, 1999; Lindgren *et al.*, 2000).

Detailed examination of the DNA regulatory elements that activate middle genes have indicated that while the MSEs of these genes are essential to their correct developmental regulation, other neighboring sequences play an important role in enhancing MSE-dependent

activation. For example, binding sites for the general transcriptional regulator ABF1 are occasionally found adjacent to MSEs and these sites have been shown in some cases to enhance MSE-dependent activation ~10-fold (Hepworth *et al.*, 1995; Ozsarac *et al.*, 1997). The close proximity of these sites to the MSEs suggests that Ndt80 and ABF1 might cooperatively bind DNA to ultimately activate transcription. This situation might be similar to the cooperative recognition of DNA by the monomeric Ig-fold transcription factor NFAT, and Fos–Jun bZIP heterodimers (L.Chen *et al.*, 1998). Cooperative interactions between NFAT and Fos–Jun allow the integration of different signaling pathways and are critical for the transcriptional regulatory events that accompany T-cell activation (Jain *et al.*, 1992, 1993).

Functional similarities between Ndt80 and p53

The significant structural and DNA-binding similarities between Ndt80 and the metazoan Ig-fold transcription factors suggest that these proteins might be evolutionarily related, in spite of their low level of primary sequence conservation. Indeed, there is some evidence that other Ig-fold transcription factors, especially p53, may be involved in regulation of meiosis. The *C.elegans* p53 homolog, *cep-1*, not only mediates the cellular response to DNA damage, but, like Ndt80, also regulates progression through meiosis and is essential for correct chromosome segregation during meiosis I (Derry *et al.*, 2001). While mammalian p53 is dispensable for meiosis (Gersten and Kemp, 1997), high levels of p53 expression have been found in tetraploid primary spermatocytes at the pachytene phase, which suggests a non-essential role for p53 at this stage of spermatogenesis (Schwartz *et al.*, 1993; Sjoblom and Lahdetie, 1996). Thus, the checkpoints that regulate cell cycle progression in response to two different kinds of DNA alterations—meiotic recombination and exogenous DNA damage—may in fact share an evolutionary heritage. If this is true, the mechanism by which Ndt80 responds to the recombination checkpoint may shed light on the complex mechanism of p53 activation in tumor suppression.

Materials and methods

Protein expression and purification

Full-length Ndt80 was expressed in *E.coli* as a fusion with maltose binding protein (MBP) using the expression plasmid pMAL-NDT80. A 2.4 kb genomic DNA fragment encoding the *NDT80* gene was used as a template to produce the *NDT80* open reading frame for pMAL-NDT80. The *NDT80* open reading frame was amplified with Vent DNA polymerase (New England Biolabs) and ligated between the *Bam*HI and *Sal*II sites in the expression vector pMAL-C2 (New England Biolabs). MBP–Ndt80 fusion protein was expressed and purified by maltose affinity chromatography as described previously (Chu and Herskowitz, 1998).

Fragments of Ndt80 were expressed in *E.coli* as fusion proteins with GST. For GST–Ndt80(1–340), the region of *NDT80* encoding residues 1–340 was amplified by PCR and cloned into pGEX-6P1 (Amersham BioSciences), between the *Bam*HI and *Xho*I restriction sites. For GST–Ndt80(59–340), the region encoding residues 59–340 was cloned into pGEX-KG. Both proteins were expressed in *E.coli* BL21. Fresh 2× YTA media was inoculated with 50 ml/l saturated overnight culture and grown at 30°C to an OD₆₀₀ = 1. The culture was induced with a final concentration of 0.25 mM IPTG and grown for an additional 6 h at 30°C. The cells were harvested and flash frozen in liquid nitrogen. Cells were then resuspended in PBS pH 7.3, 1 mM EDTA, 5 mM DTT, pepstatin,

leupeptin, benzamidine, PMSF and 0.5 mg/ml lysozyme and gently mixed on ice for 30 min, followed by sonication. The lysate was cleared by centrifugation and the supernatant was then mixed at room temperature with glutathione–Sepharose 4B beads for 1 h. This slurry was poured into a column and washed with 50 bed volumes of PBS containing 1 M NaCl, followed by 50 bed volumes of PBS. Fusion protein was eluted with PBS pH 8.0 containing 20 mM glutathione. Ndt80(1–340) was liberated from GST by cleavage with PreScission protease (Amersham BioSciences) at 4°C overnight, while Ndt80(59–340) was liberated from GST–Ndt80(59–340) by thrombin cleavage. Protein was further purified by cation exchange (SP–Sepharose FF) and gel-filtration (Superdex 75) chromatography. The resulting protein was concentrated to 20 mg/ml in a buffer containing 10 mM Tris pH 7.0, 100 mM NaCl, 1 mM DTT.

EMSA

All of the EMSAs were performed on chemically synthesized blunt end 20mer DNA using the MSE or MSE variant as shown, with flanking sequences matching the *SPS4* MSE flanking sequences. Duplex DNA was radiolabeled using T4 kinase and [γ -³²P]ATP. All the binding reactions contained 75 mM KCl, 10 mM HEPES pH 7.9, 11 mM MgCl₂, 50 μ M ZnSO₄, 10% glycerol, 1 mM DTT, 0.25 mM EDTA, 0.025% bromophenol blue and 0.25 nM radiolabeled MSE DNA (Hepworth *et al.*, 1998). All protein dilutions were made using 10 mM Tris pH 7.0, 100 mM NaCl, 1 mM DTT and 0.1 mg/ml BSA. Purified Ndt80(1–340) was the final component added to the binding reaction, which was subsequently equilibrated for 15 min at room temperature. The reaction was then loaded onto an 8% polyacrylamide gel pre-run at 100 V for 30 min in 0.5× TBE and run for another 2 h at 150 V. Bands were then visualized and quantified using PhosphorImaging plates and the program ImageQuant (Molecular Dynamics). The competition experiments were performed under the same conditions as above, but with the ³²P-labeled DNA and protein given a 5 min pre-incubation before competitor DNA was added. After addition of competitor, the reaction was incubated for 15 min at 20°C and loaded onto the gel. The initial EMSAs used in the deletion analysis were performed with induced, soluble cell lysates from *E.coli* transformed with the various deletion mutants in the presence of 10-fold weight excess of poly(dI–dC).

Proteolytic mapping

Flexible regions in Ndt80(1–349) were characterized by limited proteolytic mapping. Purified GST–Ndt80(1–349) was subjected to digestion with trypsin or thrombin over a 24 h time course, either alone or in the presence of a 5-fold molar excess of MSE DNA. Stable fragments were subsequently purified by cation-exchange (SP–Sepharose FF) and gel-filtration (Superdex 75) chromatography and their masses were determined by MALDI-TOF mass spectrometry. This analysis indicated that Ndt80(59–340) is the major trypsin/thrombin-resistant Ndt80 fragment in the absence of DNA, while residues 1–340 are protected when bound to the MSE.

In vivo assay of Ndt80 function

A 2.4 kb genomic clone of *NDT80* was subjected to site-directed mutagenesis to create a series of N-terminal deletions. Both the full-length and mutant open reading frames remained under the regulation of 600 bp of *NDT80* upstream sequence. Full-length and mutant genes were then inserted into the *TRP1* marked centromere plasmid YCplac22 (Gietz and Sugino, 1988). The mutant versions of *NDT80* were assayed for function by introducing them into DSY1257 α/α *ura3/ura3 leu2::hisG/leu2::hisG trp1::hisG/trp1::hisG ndt80::KAN^R/ndt80::KAN^R*; this strain is a homozygous α/α diploid of the SK1 lineage (Kane and Roth, 1974). The endogenous *NDT80* gene was deleted and replaced by a kanamycin resistance cassette derived from pFA6-kanMX2 (Wach *et al.*, 1994). The MSE-regulated reporter gene used to assay the function of the mutant versions of Ndt80 was *SPS4-LacZ* (Hepworth *et al.*, 1995), generously provided by Dr J.Segall. Yeast cells carrying *NDT80* or mutant versions of *NDT80* and reporter gene plasmids were initially grown in selective medium and were induced to initiate sporulation as described previously (Stuart and Wittenberg, 1998). After 6 h in sporulation medium, samples were recovered and assayed for β -galactosidase activity. The β -galactosidase activities were determined in triplicate and the data reported in Miller units as the mean \pm SD from three independent transformants. Sporulation was monitored by microscopic examination of cultures after 24 h in sporulation medium. Two hundred cells from each culture were scored for ascus formation.

DNA purification

Synthetic DNA oligonucleotides were purified on a Source 15Q column under denaturing conditions (10 mM NaOH), desalted with a C18 cartridge in a volatile buffer, lyophilized and resuspended in 5 mM Tris pH 7.0. DNA solutions were quantified by absorbance at 260 nm. Duplex DNA was annealed by heating to 80°C and slow cooling to room temperature in 5 mM Tris pH 7.0 and 100 mM NaCl at a final concentration of 1 mM DNA duplex.

Crystallization and data collection

Ndt80(1–340)–MSE complexes were prepared to a protein concentration of 10 mg/ml at a ratio of protein:DNA of 1:1. Crystals were grown using the hanging drop vapor diffusion method at room temperature (20°C) in conjunction with streak seeding. Reservoir solution (2 µl; 30% PEG 400, 50 mM bis-tris-propane pH 7.0, 100 mM NaCl, 50 mM CaCl₂, 1.5 mM spermine, 1 mM DTT) and 2 µl of Ndt80–MSE complex were mixed and immediately streaked with a hair dipped in streak solution and rinsed twice in the reservoir. Streak solution was prepared by an ~50-fold dilution of a drop that contained a shower of small crystals. Crystals grew to a maximal size of 400 µm in ~1 week and were harvested and frozen in reservoir solution. Data were collected at beamline 9-2 at Stanford Synchrotron Radiation Laboratory on an ADSC Q4 CCD. The crystals belong to space group *C*222₁ (*a* = 70.13 Å, *b* = 78.81 Å and *c* = 161.39 Å) with one molecule in the asymmetric unit. Data were processed with DENZO and SCALEPACK (Otwinowski and Minor, 1997).

Ndt80(59–340) crystals were grown using the hanging drop method at 20°C from 4 mg/ml protein solutions (which also contained a 20mer DNA duplex at a 1:1 molar ratio), equilibrated against a reservoir solution of 600 mM sodium acetate, 2 mM MgCl₂, 1 mM spermine, 25% PEG 4000 and 100 mM sodium acetate pH 5.5. Equal amounts of protein and reservoir solutions were mixed and crystals grew to a maximal size of 300 µm in 2 weeks. Cryoconditions are equivalent to the reservoir solution plus 15% glycerol. Data were collected at the Advanced Photon Source, BioCARS beamline 14-BM-D on an ADSC Q4 CCD and processed with DENZO and SCALEPACK.

Structure determination and refinement

The DNA-bound and unbound Ndt80 structures were each solved independently using a three-wavelength seleno-methionine MAD experiment. Selenium sites were located using SOLVE (Terwilliger and Berendzen, 1999) and phases were improved using both RESOLVE (Terwilliger, 2000) and DM (Cowtan, 1994). The unbound structure was built manually using O (Jones *et al.*, 1991), and refined against λ 3 of the MAD data set with iterative cycles of CNS refinement (Brünger *et al.*, 1998) and manual rebuilding.

The initial model of Ndt80 bound to DNA was created using ARP/wARP (Perrakis *et al.*, 1999) with further manual building with O. Automated refinement was carried out with CNS and REFMAC (Murshudov *et al.*, 1999) and protein geometry was analyzed with PROCHECK (Laskowski *et al.*, 1993). The final Ndt80–MSE model contains Ndt80 residues 33–139, 146–286, 294–335 and all of the DNA bases [although poor electron density is observed for the overhanging nucleotides T(–3) and A12']. The final Ndt80–MSE model contains 345 water molecules. The atomic coordinates of the Ndt80–MSE complex, and the free Ndt80 structure, have been deposited in the Protein Data Bank [PDB ID: 1MNN (Ndt80(1–340)–MSE), 1MN4 (Ndt80(59–340))].

Acknowledgements

We wish to thank Keith Brister and the staff at BioCARS beamline 14 (APS) and Peter Kuhn and the staff at SSRL beamline 9-2 for outstanding technical support during X-ray data collection. We also wish to thank Scott Williams, Ross Edwards, Ruth Green, Andrew MacMillan and members of the Glover laboratory for helpful discussions, critical reading of the manuscript and help with X-ray data collection. This work was supported by grants from the Canadian Institutes of Health Research (CIHR) and Alberta Heritage Foundation for Medical Research (AHFMR) to J.N.M.G. and D.S., and the Alberta Synchrotron Institute. J.S.L. was supported by an AHFMR studentship.

References

- Aravind,L. and Landsman,D. (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.*, **26**, 4413–4421.
- Becker,S., Groner,B. and Muller,C.W. (1998) Three-dimensional structure of the Stat3 β homodimer bound to DNA. *Nature*, **394**, 145–151.
- Bork,P., Holm,L. and Sander,C. (1994) The immunoglobulin fold. Structural classification, sequence patterns and common core. *J. Mol. Biol.*, **242**, 309–320.
- Bravo,J., Li,Z., Speck,N.A. and Warren,A.J. (2001) The leukemia-associated AML1 (Runx1)–CBF β complex functions as a DNA-induced molecular clamp. *Nat. Struct. Biol.*, **8**, 371–378.
- Brünger,A.T. *et al.* (1998) Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
- Chen,F.E., Huang,D.B., Chen,Y.Q. and Ghosh,G. (1998) Crystal structure of p50/p65 heterodimer of transcription factor NF- κ B bound to DNA. *Nature*, **391**, 410–413.
- Chen,L., Glover,J.N.M., Hogan,P.G., Rao,A. and Harrison,S.C. (1998) Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature*, **392**, 42–48.
- Chen,X., Vinkemeier,U., Zhao,Y., Jeruzalmi,D., Darnell,J.E., Jr and Kuriyan,J. (1998) Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell*, **93**, 827–839.
- Cho,Y., Gorina,S., Jeffrey,P.D. and Pavletich,N.P. (1994) Crystal structure of a p53 tumor suppressor–DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346–355.
- Chu,S. and Herskowitz,I. (1998) Gametogenesis in yeast is regulated by a transcriptional cascade dependent on Ndt80. *Mol. Cell*, **1**, 685–696.
- 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.
- Cowtan,K. (1994) An automated procedure for phase improvement by density modification. *Joint CCP4 ESF–EACBM Newsl. Protein Crystallogr.*, **31**, 34–38.
- Derry,W.B., Putzke,A.P. and Rothman,J.H. (2001) *Caenorhabditis elegans* p53: role in apoptosis, meiosis, and stress resistance. *Science*, **294**, 591–595.
- Free,A., Wakefield,R.L., Smith,B.O., Dryden,D.T., Barlow,P.N. and Bird,A.P. (2001) DNA recognition by the methyl-CpG binding domain of MeCP2. *J. Biol. Chem.*, **276**, 3353–3360.
- Gersten,K.M. and Kemp,C.J. (1997) Normal meiotic recombination in p53-deficient mice. *Nat. Genet.*, **17**, 378–379.
- Ghosh,G., van Duyn,G., Ghosh,S. and Sigler,P.B. (1995) Structure of NF- κ B p50 homodimer bound to a κ B site. *Nature*, **373**, 303–310.
- Gietz,R.D. and Sugino,A. (1988) New yeast–*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six base pair restriction sites. *Gene*, **74**, 527–534.
- 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.
- Holm,L. and Sander,C. (1997) DALI/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res.*, **25**, 231–234.
- Jain,J., McCaffrey,P.G., Valge-Archer,V.E. and Rao,A. (1992) Nuclear factor of activated T cells contains Fos and Jun. *Nature*, **356**, 801–804.
- Jain,J., McCaffrey,P.G., Miner,Z., Kerppola,T.K., Lambert,J.N., Verdine,G.L., Curran,T. and Rao,A. (1993) The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. *Nature*, **365**, 352–355.
- Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.
- Kane,S.M. and Roth,R. (1974) Carbohydrate metabolism during ascospore development in yeast. *J. Bacteriol.*, **118**, 8–14.
- 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.
- Lindgren,A., Bungard,D., Pierce,M., Xie,J., Verzhon,A. and Winter,E. (2000) The pachytene checkpoint in *Saccharomyces cerevisiae* requires the sum1 transcriptional repressor. *EMBO J.*, **19**, 6489–6497.
- Lydall,D., Nikolsky,Y., Bishop,D.K. and Weinert,T. (1996) A meiotic

- recombination checkpoint controlled by mitotic checkpoint genes. *Nature*, **383**, 840–843.
- Madej,T., Gibrat,J.F. and Bryant,S.H. (1995) Threading a database of protein cores. *Proteins*, **23**, 356–369.
- Muller,C.W. and Herrmann,B.G. (1997) Crystallographic structure of the T domain–DNA complex of the Brachyury transcription factor. *Nature*, **389**, 884–888.
- Muller,C.W., Rey,F.A., Sodeoka,M., Verdine,G.L. and Harrison,S.C. (1995) Structure of the NF- κ B p50 homodimer bound to DNA. *Nature*, **373**, 311–317.
- Murshudov,G.N., Vagin,A.A., Lebedev,A., Wilson,K.S. and Dodson,E.J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D*, **55**, 247–255.
- 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.
- Ohki,I., Shimotake,N., Fujita,N., Jee,J., Ikegami,T., Nakao,M. and Shirakawa,M. (2001) Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell*, **105**, 487–497.
- 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. *Methods Enzymol.*, **276**, 307–325.
- 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.
- Perrakis,A., Morris,R. and Lamzin,V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.*, **6**, 458–463.
- Pijnappel,W.W. *et al.* (2001) The *S.cerevisiae* SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev.*, **15**, 2991–3004.
- Roeder,G.S. (1997) Meiotic chromosomes: it takes two to tango. *Genes Dev.*, **11**, 2600–2621.
- Rooman,M., Lievin,J., Buisine,E. and Wintjens,R. (2002) Cation- ϕ /H-bond stair motifs at protein–DNA interfaces. *J. Mol. Biol.*, **319**, 67–76.
- Rudolph,M.J. and Gergen,J.P. (2001) DNA-binding by Ig-fold proteins. *Nat. Struct. Biol.*, **8**, 384–386.
- Schwartz,D., Goldfinger,N. and Rotter,V. (1993) Expression of p53 protein in spermatogenesis is confined to the tetraploid pachytene primary spermatocytes. *Oncogene*, **8**, 1487–1494.
- Sjoberg,T. and Lahdetie,J. (1996) Expression of p53 in normal and γ -irradiated rat testis suggests a role for p53 in meiotic recombination and repair. *Oncogene*, **12**, 2499–2505.
- Stuart,D. and Wittenberg,C. (1998) CLB5 and CLB6 are required for premeiotic DNA replication and activation of the meiotic S/M checkpoint. *Genes Dev.*, **12**, 2698–2710.
- Suzuki,M., Yagi,N. and Finch,J.T. (1996) Role of base–backbone and base–base interactions in alternating DNA conformations. *FEBS Lett.*, **379**, 148–152.
- Tahirov,T.H. *et al.* (2001) Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBF β . *Cell*, **104**, 755–767.
- Terwilliger,T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D*, **56**, 965–972.
- Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D*, **55**, 849–861.
- 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.
- Wach,A., Brachat,A., Pohlmann,R. and Philippsen,P. (1994) New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast*, **10**, 1793–1808.
- Wintjens,R., Lievin,J., Rooman,M. and Buisine,E. (2000) Contribution of cation- ϕ interactions to the stability of protein–DNA complexes. *J. Mol. Biol.*, **302**, 395–410.
- 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.

Received July 10, 2002; revised and accepted September 6, 2002