doi:10.1006/jmbi.2000.3606 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **297**, 947–959

JMB

Crystallization of the Yeast MATα2/MCM1/DNA Ternary Complex: General Methods and Principles for Protein/DNA Cocrystallization

Song Tan, Yvonne Hunziker, Luca Pellegrini and Timothy J. Richmond*

ETH-Zürich, Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg CH-8093 Zürich, Switzerland

We describe our efforts to crystallize binary MCM1/DNA and ternary MATα2/MCM1/DNA complexes, including the unsuccessful attempts to crystallize MCM1/DNA complexes and the successful design of DNA crystal packing that resulted in high-resolution crystals of the MATα2/MCM1/DNA complex. We detail general procedures useful for preparing protein/DNA cocrystals, including improved methods for producing and purifying DNA-binding proteins and DNA fragments, for purifying protein/DNA complexes, and for controlling pH conditions during crystallization. We also describe the rational design of DNA for protein/DNA cocrystallization attempts, based on our analysis of how straight and bent DNA with single base-pair overhangs can pack end-to-end in a crystal.

Keywords: protein/DNA cocrystallization; X-ray crystallography; yeast mating type transcription factors; protein purification; MATα2 transcription factor

**Corresponding author*

Introduction

Our understanding of the molecular interactions between DNA-binding proteins and DNA has benefitted greatly from X-ray and NMR structure determination of protein/DNA complexes and the resulting detailed views of the interaction surfaces. X-ray crystallography is the method of choice for protein/DNA complexes with molecular masses of greater than 30 kDa, but crystallizing the complex remains the rate-limiting step for many structural determinations. The difficulty of obtaining diffraction-quality macromolecular crystals can usually be traced to the need for large quantities of high-purity material and the empirical nature of crystallization. For protein/DNA cocrystals, both protein and DNA components must be overexpressed or synthesized, and the resulting protein/DNA complex then characterized and purified. Optimization of each of these steps is often required to produce high-resolution protein/DNA crystals.

The precise choice of DNA fragments can be critical to the formation of protein/DNA cocrystals because interactions between DNA molecules often generate crystal packing contacts. The two predominant strategies for selecting DNA fragments are the empirical searches for fragments of different lengths and ends (Anderson *et al.*, 1984), and the rational design of DNA fragments that pack in defined ways in the crystal (Jordan *et al.*, 1985). In many of the protein/DNA cocrystals solved to date, the DNA packs end-to-end as a pseudo-continuous helix. Anderson *et al.* (1984) noted that DNA fragments which pack end-to-end to match the helical repeat of DNA would be appropriate choices for cocrystallization attempts. Thus, a 21 bp DNA fragment with two turns of a 10.5 bp/turn helix would be a likely candidate under this scheme, as would a 14 bp DNA fragment that, when repeated three times, produces 42 bp or four turns of a 10.5 bp/turn helix. In fact, of the 32 protein/DNA complexes in the RCSB Protein Data Bank containing DNA lengths of 20 or 21 bp as of November 1999, all but one contained zero, one or two complementary overhanging bases at the two ends, and 28 of these produced pseudocontinuous

Present addresses: S. Tan, 108 Althouse Laboratory, Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA 16802, USA; L. Pellegrini, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK.

E-mail address of the corresponding author: richmond@mol.biol.ethz.ch

0022-2836/00/040947–13 $35.00/0

DNA helices through the crystal *via* these complementary ends.

Most protein/DNA cocrystallization attempts therefore involve screening different DNA substrates, a process greatly helped by the relative ease of synthesizing and purifying milligram quantities of oligonucleotides. However, synthetic oligonucleotides are limited to sizes of 30 bases or less for the preparative-scale synthesis yields and purity needed for structural work. For this reason, at least three protein/DNA structures (CAP/DNA, γδ-resolvase/DNA and IHF/DNA) have used overlapping complementary oligonucleotides to create longer, nicked DNA fragments for crystallization (Rice *et al*., 1996; Schultz *et al*., 1990; Yang & Steitz, 1995). An alternative method to produce large quantities of crystallization-quality DNA is the isolation of restriction fragments from plasmids grown in *Escherichia coli*, as pioneered in studies of the nucleosome (Richmond *et al*., 1988). A significant advantage of using restriction fragments is the ability to isolate homogeneous DNA fragments without size limitations, an important consideration for crystallizing binary or higher order complexes containing large DNA binding sites. However, it is usually necessary to construct plasmids with multiple copies of the desired insert to obtain the milligram quantities required for crystallization, and each insert in such a plasmid must be in the same orientation to avoid unstable palindromic sequences (Gupta *et al*., 1983).

Here, we detail the steps taken to produce high-resolution crystals of a ternary complex of the yeast transcription factors MCM1 and MATα2 (abbreviation, α2) bound to DNA, a complex required for the repression of mating type genes. The structure of this complex has been described elsewhere (Tan & Richmond, 1998). Our unsuccessful attempts to crystallize binary MCM1/DNA complexes and full-length α2/MCM1/DNA complexes led to a directed strategy where we optimized both protein and DNA fragments for crystallization. Our work also involved improving protein purification methods, and developing schemes for producing large quantities of defined sequence DNA fragments in *E. coli*. In addition, we present a general framework for the rational design of DNA fragments lengths and ends to guide future protein/DNA cocrystallization attempts.

Results

Improved use of heparin-Sepharose to purify DNA-binding proteins

We and others have found that many DNA-binding proteins can be purified from an *E. coli* crude extract using heparin-Sepharose. The crude extract prepared by lysis of the overexpressing cells and clarification by centrifugation can be loaded onto the resin without prior removal of nucleic acid by protamine sulfate or polyethyleneimine. One novel step we have added is washing with buffer containing moderate concentrations of urea in addition to salt washes. We reasoned that many of the contaminants eluting with the desired protein were binding non-specifically *via* hydrophobic interactions to the resin, or proteins bound to the resin. Reducing this hydrophobic interaction with urea at concentrations insufficient to denature the desired protein (2 to 3 M) removes many contaminating proteins. Figure 1 shows that significant amounts of contaminating protein can be eluted with buffer containing 3 M urea with only moderate losses of the desired MCM1(1-100), resulting in purer protein eluting from the column during the gradient elution.

Isolation of STE2 35mer operator DNA

We created plasmids that contain multiple tandem copies of a STE2 35mer, each arranged in a head-to-tail orientation to reduce stability problems encountered with palindromic repeats. Our attempts to create a plasmid with multiple inserts by directly subcloning a synthesized DNA fragment self-ligated in a fixed orientation produced a maximum of four inserts in a plasmid. We have therefore implemented a procedure, which was suggested by Richmond *et al.* (1988), to successively double the number of tandem inserts in a plasmid. This procedure relies on a pair of restriction enzymes which produce compatible ends (*Sal*I and *Xho*I in this case). The insert was subcloned into a plasmid vector such that it was flanked by *Sal*I and *Xho*I sites. As shown in Figure 2(a), isolation of the appropriate vector fragment (*Eco*RI-*Sal*I) and insert fragment (*Eco*RI-*Xho*I) and subsequent ligation results in the doubling of the insert with a short linker in between. When ligated, the compatible ends of the *Sal*I/*Xho*I pair produce a site that is recognized by neither restriction enzyme. Thus, the insert resulting from the doubling procedure is once again flanked by unique *Sal*I and *Xho*I sites, and is suitable for additional rounds of doubling. Other pairs of restriction enzymes producing compatible ends can be used, such as *Bam*HI and *Bgl*II, with the only restriction that neither the insert nor the vector fragments contain such sites. We have used this doubling procedure to produce pST17, a plasmid with 48 copies of a 35 bp fragment containing the nearly symmetric STE2 31 bp UAS element (Figure 2(b)) as well as pST5, a plasmid containing 40 copies of an 84 bp fragment used to produce symmetric 146 bp DNA fragments for studies of the nucleosome core particle (Flaus *et al*., 1996). A similar procedure for generating recombinant nucleosomal DNA, which substitutes less-efficient blunt end ligation for our compatible-end cloning strategy, has been described (Palmer *et al*., 1995).

The 35mer insert in the pST17 plasmid was released with *Eco*RV digestion and purified from the 2700 bp vector fragment and the 11 bp linker (which includes the destroyed *Sal*I/*Xho*I site) by a combination of PEG precipitation and anion-

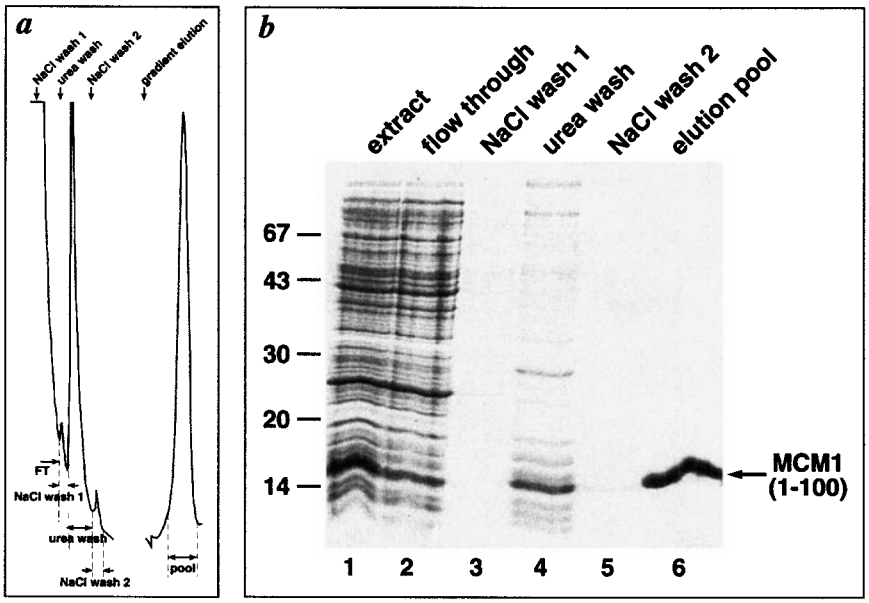

Figure 1. Heparin-Sepharose purification of MCM1(1-100) and use of a moderate concentration urea wash to remove non-specifically bound proteins. (a) UV absorption trace of eluent from heparin-Sepharose column monitored at 276 nm. Both NaCl washes contained 300 mM NaCl, while the urea wash contained an additional 3 M urea. The remaining protein was eluted with a linear gradient from 300 mM to 1500 mM NaCl. Washing and elution steps are indicated above the trace, while pooled samples are indicated below (FT, flow-through). (b) Coomassie-stained SDS-PAGE gel of heparin-Sepharose fractions. The urea wash contains significant amounts of contaminating proteins that would otherwise elute during the gradient elution (data not shown). Note that the gradient elution pool was also passed through a Matrex Blue A column which removes additional contaminants. Positions of molecular mass markers are indicated on the left.

exchange HPLC (Figure 3). We routinely obtain 100 to 120 mg of pST13 plasmid from 12 l of HB101/pST13, which yields 20 to 25 mg of purified STE2 35mer. Although the pST17 plasmid contains twice as many STE2 35mer inserts as pST13, we obtained similar quantities of insert from SURE/pST17 as from HB101/pST13 because of the reduced yield of plasmid from the SURE strain.

Crystallization of MCM1/DNA complexes

Despite multiple attempts, we were unable to crystallize full-length MCM1 alone or complexed with DNA. DNA fragments tested for the cocrystallization studies were the STE2 35mer, and symmetric P-box oligonucleotides of length 16, 17 and 18 bases (STO85, -86, -87, respectively). Since the C-terminal two-thirds of MCM1 contains an unusual sequence composition, further crystallization trials were directed at the MCM1(1-100) deletion, which is sufficient for DNA-binding *in vitro* and mating-type function *in vivo* (Bruhn *et al*., 1992; Christ & Tye, 1991; Primig *et al*., 1991). We examined 85 different DNA fragments in three sequence categories: symmetric P-box sequences, **a**-specific sequences such as STE2 UAS elements, and α-specific sequences such as STE3 UAS elements (Table 1). Furthermore, we incorporated different ends on the DNA fragments to encourage end-to-end or other packing of DNA in the crystal. In addition to blunt ends, single base-pairs overhangs were included to allow Watson-Crick base-pairing or triple strand formation. Only small crystals unsuitable for diffraction studies were obtained with MCM1(1-100)/P-box complexes, whereas no crystals were obtained for MCM1(1-100)/STE2 UAS complexes.

Attempts to crystallize MCM1(1-100)/STE3 complexes met with limited success. All the MCM(1-100) cocrystals that we obtained contained variants of a STE3 UAS 18 bp fragment. MCM1(1-100)/(STO60/61) blunt end DNA cocrystals diffract to about 10 Å, with characteristic streaky spots. When single-base overhangs are added compatible with triple strand formation at the ends (e.g. STO138/139), the resulting crystals diffracted to about 8 Å without the streaky spots observed with blunt ended DNA cocrystals. Considerable effort was needed to acquire MCM1(1-100) crystals of sufficient size for diffraction studies, since showers of small crystals were initially obtained. The two successful modifications were preparative electrophoresis to purify the protein/DNA complex and use of a slow pH gradient during crystallization.

Our experiments show that preparative electrophoresis purified MCM1(1-100)/(STO138/139) produces larger crystals by reducing the number of nucleation events, presumably by removing aggregates that might be formed by excess protein. Attempts to control nucleation with microseeding were not successful. Since many DNA-binding proteins are positively charged overall, and since pro-

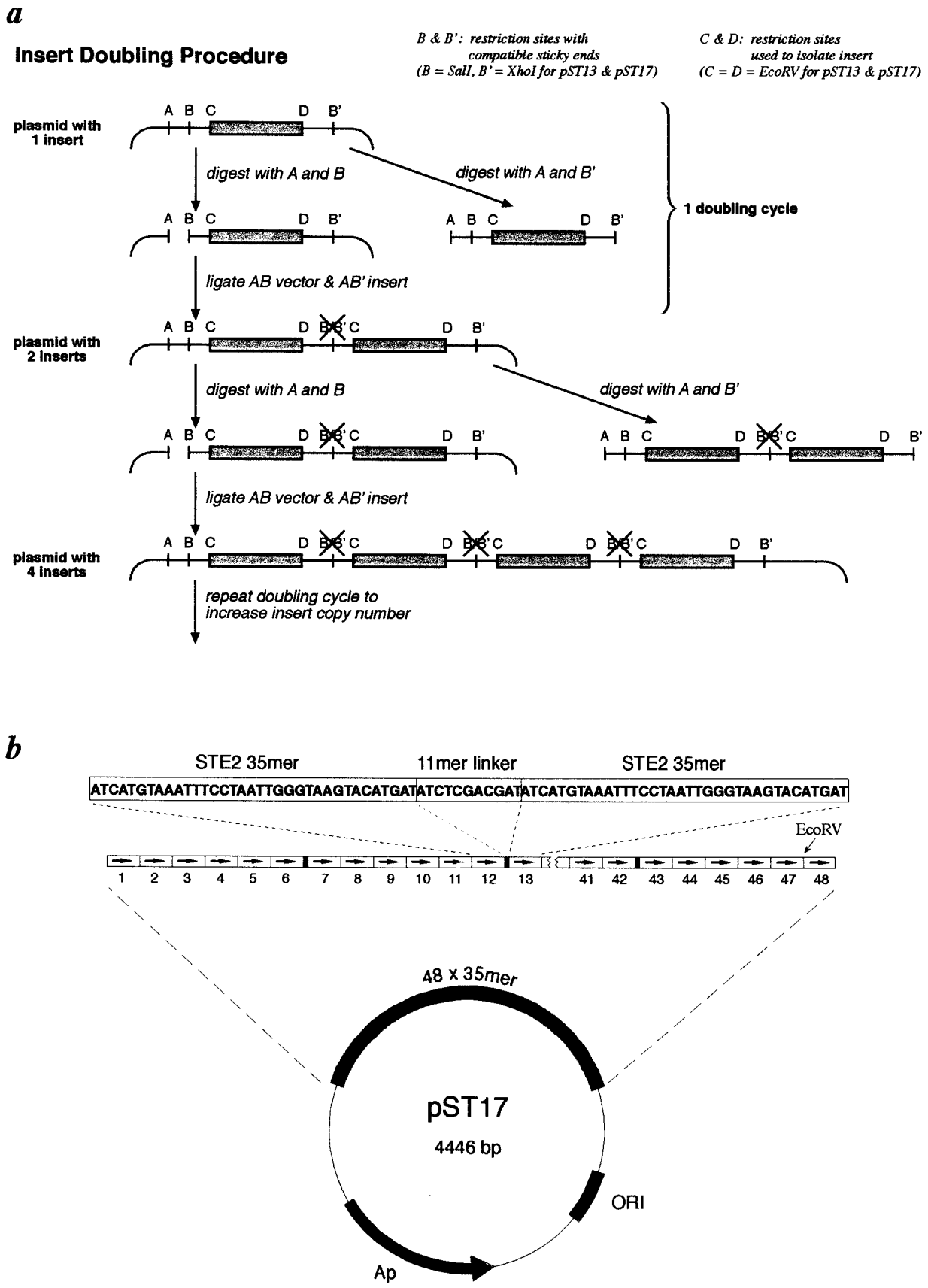

Figure 2. Multi-copy insert plasmids. (a) Schematic of insert doubling procedure. The desired insert defined by restriction sites C and D, and flanked by the restriction sites B and B′ with compatible ends to each other, can be duplicated in a plasmid. This doubling procedure can be reiterated to further increase the number of insets in the

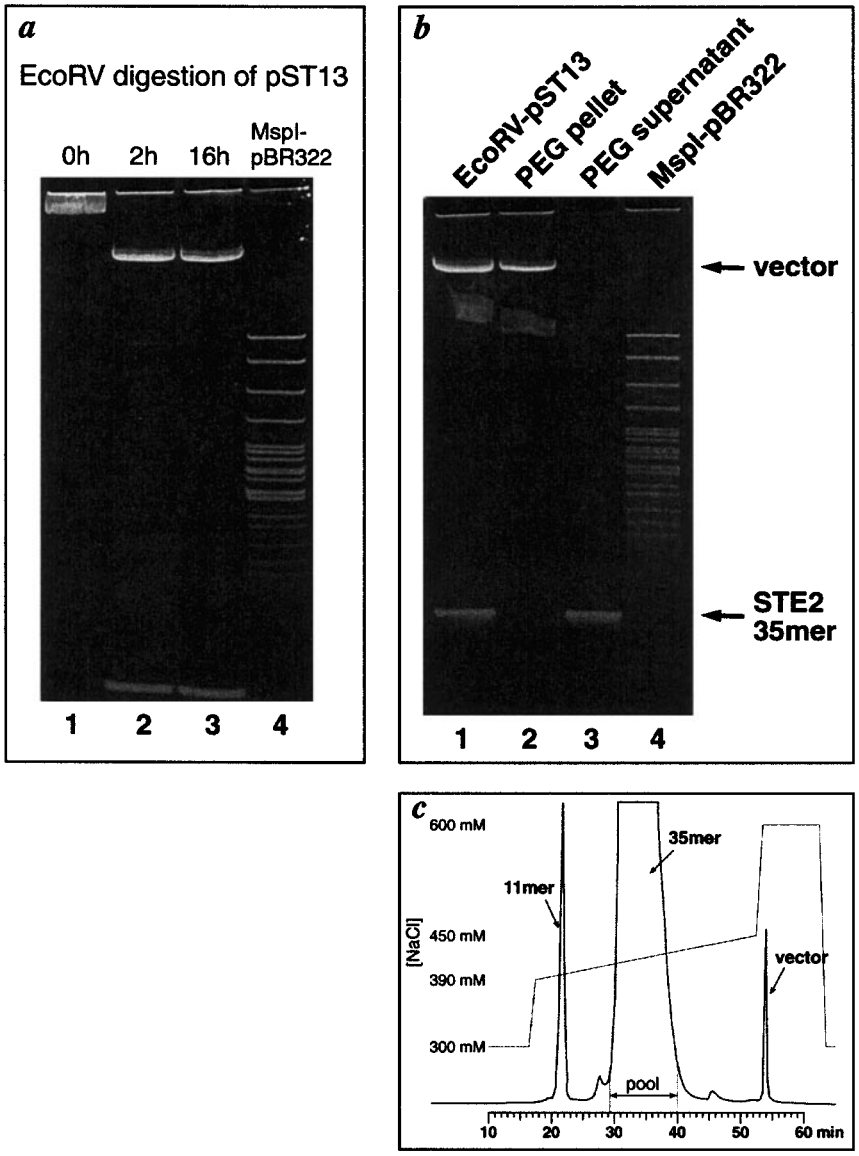

Figure 3. Isolation of the STE2 35mer from the multiple copy insert plasmid pST13. (a) Ethidium bromide stained polyacrylamide gel of *Eco*RV digestion of pST13 to release STE2 35mer insert. The undigested plasmid shown in lane 1 does not enter the gel. Digestion was complete within two hours, since no partial digestion products were detected after two hours (lane 2) or 16 hours (lane 3). *Msp*I-digested pBR322 molecular mass markers are shown in lane 4. (b) Ethidium bromide stained PAGE of polyethylene glycol precipitation of *Eco*RV digestion products. The *Eco*RV digestion products are shown in lane 1, the PEG precipitate in lane 2 and the PEG supernatant in lane 3. Positions of the 2700 bp vector and 35 bp insert fragments are indicated. Although the 11 bp linker might be visible on the gel shown, it is rarely detected by ethidium bromide staining because of the relative scarcity of this short fragment. (c) Purification of STE2 35mer from PEG supernatant by anion-exchange HPLC. The absorption trace (monitored at 260 nm) of the HPLC run is shown, together with the NaCl gradient used (corrected for void volume of column and tubing) and the identity of the major peaks. The STE2 35mer peak saturated the detector flow-cell with a maximum absorbance of 2.0.

Figure 2. (*Legend continued*)

plasmid, since each plasmid will contain a unique B and B′ site. Four rounds of the doubling procedure were used to increase the insert copy number of six in pST11, to 48 copies in pST17. (b) Schematic plasmid containing 48 copies of STE2 35mer. The arrangement of inserts in the plasmid shows that tandem blocks of six contiguous repeats are separated by an 11 bp linker, which contains the two half-sites for *Eco*RV as well as the sequence CTCGAG which results from the ligation of *Sal*I and *Xho*I ends. Sequences for the STE2 35mer and the 11 bp linker are presented, with the *Eco*RV sites underlined.

Table 1. Summary of oligonucleotides used for cocrystallization attemps with MCM1

No.	Name	Type	Size	Ends	Direction	MCM1(1-100) xtals	Diffraction (Å)
1	STO85	P-box	16	Blunt	↔	-	
2	STO86	P-box	17	5′T	↔	-	
3	STO87	P-box	18	Blunt	↔	-	
4	STO88	P-box	19	5′T	↔	-	
5	STO89	P-box	20	Blunt	↔	-	
6	STO90	P-box	21	5′T	↔	-	
7	STO106	P-box	22	Blunt	↔	-	
8	STO132	P-box	17	C ~ GC	↔	Needles	
9	STO133	P-box	19	C ~ GC	↔	-	
10	STO134	P-box	21	G ~ GC	↔	-	
11	STO135	P-box	23	G ~ GC	↔	-	
12	STO136	P-box	25	C ~ GC	↔	-	
13	STO137	P-box	27	C ~ GC	↔	-	
14	STO155/156	P-box	33	G ~ GC	↔	-	
15	STO188	P-box	21	C ~ GC	↔	Tiny needles	
16	STO189	P-box	23	G ~ GC	↔	Tiny needles	
17	STO148	P-box	10	Blunt	↔	-	
18	STO149	P-box	12	Blunt	↔	Thin plates	
19	STO150	P-box	13	C ~ GC	↔	Thin plates	
20	STO151	P-box	13	G ~ GC	↔	Thin plates	
21	STO152	P-box	14	Blunt	↔	Needles	
22	STO153	P-box	15	C ~ GC	↔	-	
23	STO154	P-box	15	G ~ GC	↔	-	
24	STO91/92	STE2	20	Blunt	↔	-	
25	STO96/97	STE2	20	WC	→→	-	
26	STO98/99	STE2	21	WC	→→	-	
27	STO101/102	STE2	22	WC	→→	-	
28	STO103/104	STE2	26	WC	→→	-	
29	STO105/104	STE2	26	Blunt	↔	-	
30	STO123/124	STE2	31	Blunt	↔	-	
31	STO217/218	STE2	18	C ~ GC,T ~ AT	→→	-	
32	STO219/220	STE2	20	C ~ GC,T ~ AT	→→	-	
33	STO221/222	STE2	22	C ~ GC,T ~ AT	→→	-	
34	STO234/235	STE2	17	WC	→→	-	
35	STO236/237	STE2	18	WC	→→	-	
36	STO238/239	STE2	19	WC	→→	-	
37	STO250/251	STE2	21	WC	→→	-	
38	STO252/253	STE2	22	WC	→→	-	
39	STO240/241	STE2	23	WC	→→	-	
40	STO242/243	STE2	24	WC	→→	-	
41	STO244/245	STE2	25	WC	→→	-	
42	STO246/247	STE2	26	WC	→→	-	
43	STO248/249	STE2	27	WC	→→	-	
44	STO60/61	STE3	18	Blunt	↔	xtals	10
45	STO93/94	STE3	19	WC	→→	-	
46	STO95	STE3	20	Blunt	↔	-	
47	STO61/100	STE3	18	WC	→→	-	
48	STO93/107	STE3	19	5′T	→→	-	
49	STO108/109	STE3	20	Blunt	↔	-	
50	STO112/113	MFα2	18	Blunt	↔	Small xtals	
51	STO114/115	MFα1A	18	Blunt	↔	Small xtals	
52	STO117/118	MFα1B	18	Blunt	↔	Small xtals	
53	STO119/120	MFα1C	18	Blunt	↔	xtals	No diffraction
54	STO121/122	STE3	18	5′T	↔	-	
55	STO129/130	STE3	19	C ~ GC,T ~ AT	→→	-	
56	STO138/139	STE3	19	C ~ GC	↔	xtals	8
57	STO140/141	STE3	20	C ~ GC	↔	-	
58	STO142/143	STE3	21	G ~ GC	↔	-	
59	STO144/145	STE3	22	G ~ GC	↔	-	
60	STO146/147	STE3	23	G ~ GC	↔	-	
61	STO157/158	STE3	19	T ~ AT	↔	-	
62	STO159/160	STE3	19	G ~ GC	↔	-	
63	STO161-4	STE3	18	Blunt, nicked	↔	-	
64	STO167/168	STE3	18	Blunt	↔	-	
65	STO169/170	STE3	18	Blunt	↔	-	
66	STO190/191	MFα2	19	G ~ GC	↔	Small xtals	
67	STO192/193	MFα1A	19	G ~ GC	↔	-	
68	STO194/195	MFα1B	19	G ~ GC	↔	Small xtals	
69	STO196/197	MFα1C	19	G ~ GC	↔	-	
70	STO198/199	STE3	19	C ~ GC	↔	xtals	10
71	STO20/201	STE3	19	C ~ GC	↔	xtals	15
72	STO202/203	STE3	19	C ~ GC	↔	xtals	No diffraction

Table 1. (*Continued*).

73	STO204/205	STE3	19	C ~ GC	↔	xtals	10
74	STO206/207	STE3	19	C ~ GC	↔	xtals	20
75	STO208/209	STE3	19	C ~ GC	↔	xtals	10
76	STO213/214	STE3	19	C ~ GC,T ~ AT	→→	-	
77	STO215/216	STE3	19	C ~ GC,T ~ AT	→→	-	
78	STO224/225	STE3	24	C ~ GC	↔		
79	STO226/227	STE3	24	C ~ GC	↔	-	
80	STO228/229	STE3	25	C ~ GC	↔	-	
81	STO230/231	STE3	19	C ~ GC,T ~ AT	→←	xtals	8
82	STO232/233	STE3	19	C ~ GC,T ~ AT	→←	Small xtals	

Type indicates the sequence (respective promoter or P-box consensus sequence).
The length of each of the two strands that comprise the cocrystallization DNA is given in the size column.
The type of DNA ends is indicated as WC for Watson-Crick overhang, C ~ GC for C ~ GC triple strand overhang, T ~ AT for T ~ AT triple strand overhang, 5′N for single N base overhang or blunt.
The expected direction of the cocrystallization DNA if packed end to end is provided as →→ if adjacent DNA fragments pack head-to-tail, →← if they pack head-to-head and ↔ if they can pack in either direction.
Results of crystallization trials of MCM1 (1-100)/DNA complexes are provided in the MCM1(1-100) xtals column.
A dash indicates that no crystals were observed.
The highest resolution for which diffraction was detected is presented in the diffraction column.

tein/DNA complexes often migrate significantly slower then DNA alone, purification of the protein/DNA complex from excess DNA or protein is usually easy to achieve, requiring little optimization beyond the concentration of acrylamide in the gel. The preparative electrophoresis system that we used offers a further benefit that the elution buffer can be different from the electrophoresis buffer, enabling the sample to be concentrated immediately without additional dialysis. We also used preparative electrophoresis to purify the homologous SRF/DNA complex prior to crystallization and subsequent structure determination (Pellegrini *et al*., 1995). Other examples of preparative electrophoresis to complexes prior to crystallization include the interferon γ/interferon γ receptor complex (Fountoulakis *et al*., 1993) and the nucleosome core particle (Harp *et al*., 1995; Luger *et al*., 1997a,b).

Our observation that the MCM1(1-100)(STO138/139) crystals were sensitive to variations of 0.1 pH unit prompted us to use a volatile buffer such as acetate to change the pH of the crystallization drop over several days. This acute sensitivity to pH conditions, which was also observed during crystallization of the glucocorticoid receptor/DNA complex (B. Luisi, personal communication), may reflect the requirement of a protonated cytosine base to form the Hoogstein base-pairing at the triple strand end. The use of a pH gradient during crystallization significantly increased the reproducibility of growing suitably large crystals, probably because the pH gradient allows a range of pH to be sampled during crystallization. This increases the chance of the sample encountering the appropriate conditions for nucleation and/or growth as compared to a discrete set of pH conditions. The pH in the hanging drop will gradually approach the pH of the reservoir, with the time required for the change determined predominantly by the volatility and concentration of the volatile buffer, and the temperature of the system.

We also tried to improve the diffraction of the MCM1/DNA cocrystals by modifying both the protein and DNA components. The two MCM1 deletions examined were MCM1(1-96) and MCM1(16-98). MCM1(1-96) produced similar cocrystals with STO138/139 to MCM1(1-100) under the same crystallization conditions, whereas MCM1(16-98) required different conditions for cocrystal formation. However, no improvement in diffraction quality was observed with these polypeptide deletion variants. Varying the DNA length eliminated crystal formation, although changes in DNA sequence could be tolerated and cocrystals were obtained using other α-specific sequences such as MFα1B. Despite these and other changes such as different triple strand ends (Table 1) and use of 5′ phosphorylated oligonucleotides (data not shown), only similar or worse diffraction quality was obtained as compared to the MCM1(1-100)/STO138/139 crystals.

Crystallization of α2/MCM1/DNA complexes

We attempted crystallization of full-length α2 in a ternary complex with several MCM1 deletions and the recombinant STE2 35mer. Each of these ternary complexes containing full-length MCM1, MCM1(1-166), MCM1(1-100) or MCM1(16-98) was purified by gel-filtration chromatography prior to crystallization trials (data not shown). However, these crystallization trials did not produce any crystals, despite the use of different precipitating agents under a variety of conditions.

Our unsuccessful attempts to obtain diffraction-quality crystals of MCM1/DNA or full-length α2/MCM1/DNA complexes led us to develop a rational design to crystallize a complex containing α2, MCM1 and DNA. Since we were most interested in characterizing those interactions responsible for formation of the α2/MCM1/DNA ternary complex, we decided to eliminate polypeptide segments not required for this interaction, but we retained segments if they contributed to stability or

solubility. We therefore selected MCM1(1-100), since it contains the conserved DNA-binding domain that is sufficient for interaction with α2, as well as the first 18 amino acid residues that improve solubility of the polypeptide (unpublished results). The α2(103-189) polypeptide selected eliminates the N-terminal domain as well as the short C-terminal tail responsible for interaction with MAT**a**1, but retains the homeodomain and the N-terminal extension that is required and sufficient for interaction with MCM1 (Vershon & Johnson, 1993).

Since we anticipated that MCM1 would bend DNA (based on the strong homology to SRF, our structure of the SRF/DNA complex (Pellegrini *et al*., 1995) and circular permutation studies (Acton *et al*., 1997)), we chose DNA fragments that could pack end-to-end in a crystal whether or not the DNA was bent. This was achieved by using DNA lengths equivalent to half-integral repeat of the DNA helix pitch (25-27 bp) and by incorporating appropriate DNA ends to allow end-to-end packing. Use of a 25-27 bp DNA fragment precludes two α2 binding sites flanking the central MCM1 site found in a-specific UAS elements. However, since we find that α2(103-189) does not bind cooperatively to the STE2 35mer (data not shown), we chose to use a truncated UAS element containing only one α2 binding site (Figure 4). We obtained ternary complex crystals for five of the ten DNA fragments examined (all produced by phosphoramidite chemical synthesis), including a STE6 26mer (STO577/578), for which the ternary complex crystal diffracts to 2.25 Å. α2(103-189)/MCM1(1-100)/(STO577/578) crystallizes in space group $P2_12_12_1$ with cell dimensions 70.6 Å × 72.6 Å × 150.7 Å, where the longest dimension is consistent with a bent 26 bp DNA fragment packed end-to-end along this long axis.

Discussion and Conclusion

Our success at crystallization of the ternary α2/MCM1/DNA complex was dependent on selecting the right protein and DNA components. Regions of MCM1 and α2 that were not involved in the formation of the ternary complex were eliminated, although we did retain the N-terminal 17 amino acids residues of MCM1 that increased the solubility of the protein. By removing polypeptide segments not involved in the function we were interested in investigating, we reduced the likelihood of unstructured regions which could interfere with crystallization contacts.

Our choice of cocrystallization DNA fragment was based on an analysis of how DNA can pack in a crystal. Previous analyses of possible DNA crystal packing arrangements for protein/DNA complexes were limited to straight DNA with Watson-Crick base-pairing ends. We wish to extend this crystal packing analysis to include bent DNA and the use of triple strand ends. As shown in Figure 5(a), the integral repeat DNA packing by Anderson *et al*. (1984) applies when the DNA helix is not distorted by intrinsic or protein-induced bending. An integral repeat of bent DNA will form a superhelix when packed end-to-end (Figure 5(b)), as was observed in the a1/α2/DNA $P6^1$ cocrystal with six 21mer monomers in each superhelical repeat (Li *et al*., 1995b). However, crystals containing one superhelical repeat and reasonable unit cell dimensions require an integral number of DNA monomers per superhelical turn, which constrains the DNA bend angle to discrete values. In fact, the 60° DNA bend angle in the al/α2/DNA crystal structure is less than the 100° bending estimated by solution methods (Jin *et al*., 1995; Li *et al*., 1995a; Smith *et al*., 1995). Conversely, it is possible that the integral repeat DNA packing in many of the other protein/DNA crystals may have favored relatively straight DNA. Such artificial constraints on the structure of DNA in a protein/DNA crystal can be avoided if a half-integral repeat of DNA such as 16 or 26 bp is used (assuming a *B*-form conformation), since the DNA can then pack end-to-end in a zig-zag fashion to produce a 2_1 screw axis. In this scheme, no restrictions are placed on the DNA bend angle, since any bend in one DNA

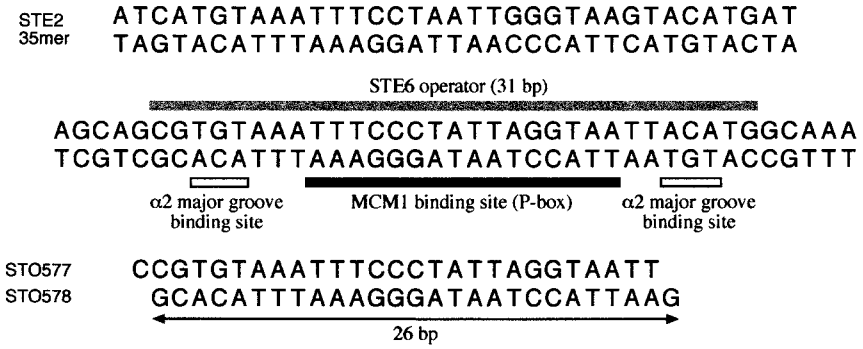

Figure 4. Sequence of STE6 operator and synthetic DNA fragment successfully used to crystallize α2/MCM1/DNA complex. The 31 bp STE6 operator with flanking sequence is shown with the location of the locations of the consensus P-box MCM1 binding site and the flanking α2 major groove binding sites. The sequences of the two synthetic oligonucleotides, STO577 and STO578, are provided below, while the sequence of the STE2 35mer is provided above for comparison.

monomer is compensated by the next monomer (Figure 5(c)).

Although Watson-Crick hydrogen bonds for promoting end-to-end packing of DNA have been heavily used in protein/DNA cocrystals, triple strand base-pairings provide an important alternative. The glucocorticoid receptor/DNA (Luisi *et al*., 1991), SRF/DNA (Pellegrini *et al*., 1995) and CAP/DNA (Schultz *et al*., 1990) complexes showed that C or G can form a triple strand in the major groove of a G·C base-pair for an alternate end-to-end packing (Figure 5(d)-(i)). Both the C^+-GC (Hoogstein interactions, also true for A-TA) and G-GC ends result in a disjointed DNA helix as compared to the continuous helix with Watson-Crick ends (Figure 5(f) and (g)). The disjointed helix with C^+-GC triple strand ends involves an angular displacement equivalent to backing up the helix by 2 bp in addition to a translational displacement (Figure 5(g) and (i)). Thus the effective DNA helical twist repeat for DNA fragments forming C^+-GC (or A-TA) triple strands ends is 2 bp less than the actual repeat length. Both the glucocorticoid receptor/DNA and SRF/DNA complexes used 18 bp DNA fragments with a single base overhang. This corresponds to an effective DNA helical twist repeat of 16 bp, which is a half-integral helical twist repeat appropriate for structures with both straight (glucocorticoid/DNA) and bent (SRF/DNA) DNA (Figure 5(e)). Similarly, a half-integral helical twist of DNA was also present in the CAP/bent DNA structure which used the equivalent of a 30 bp fragment with single base overhangs. The disjointed helix in the CAP/DNA structure with the G-GC triple strand ends backs up the helix by 3 bp, so that the 30 bp DNA repeat corresponds to a 27 bp effective helical twist repeat or approximately 2.5 helical repeats (Figure 5(d), (f) and (h)).

The major benefit of triple-stranded ends is to increase the range of DNA lengths compatible with end-to-end packing in a crystal. Furthermore, the distance between the binding sites increases by 2 bp for C^+-GC and G-GC triple strand ends compared to Watson-Crick base-pairing ends. For example, the repeat length of the 19mer with C^+-GC ends used in the SRF/DNA complex is 18 bp, since the triple strand base-pairings at each end result in the net loss of 1 bp of repeat length. As described above, this corresponds to a 16 bp helix repeat length or one and a half turns of the DNA helix. In contrast, the equivalent DNA fragment using Watson-Crick base-pairing would be a 16mer with complementary overhanging ends. Such a fragment would possess both 16 bp repeat length, as well as 16 bp helical twist repeat. This additional 2 bp spacing of repeat length afforded by the triple-stranded ends (as compared to the Watson-Crick ends) might be useful, since it reduces the possibility of steric hindrance between protein components on adjacent DNA fragments. An efficient strategy for protein/DNA cocrystallization is to vary both DNA length, to allow for DNA helical repeats other than *B*-form 10.5 bp/turn, as well as DNA ends, to accommodate Watson-Crick or triple-stranded base-pairing ends. The ability to vary such parameters is likely to increase the variety of crystal packing possible and therefore the likelihood of obtaining crystals.

We have used these improved methods for protein and DNA expression and purification, design of protein/DNA complexes and crystallization to produce high-resolution α2/MCM1/DNA ternary complex crystals. These crystals have allowed us to determine the structure of the complex at 2.25 Å, which explains how α2 and MCM1 bind cooperatively to DNA, as well as documenting an unusual structural flexibility of α2 (Tan & Richmond, 1998). The structure also shows that the DNA in the crystal does pack in a zig-zag manner as predicted (Figure 5(c)). However, this designed crystal packing in one dimension was accompanied by other fortuitous crystal packing, including the incorporation of a second copy of α2 which provides perpendicular packing interactions by bridging between adjacent complexes in the crystal (see Figure 1(c) by Tan & Richmond, 1998). Thus, our designed DNA crystal contacts may have improved the chances of obtaining crystals, but we are still unable to predict or design all crystal interactions.

Materials and Methods

Overexpression and purification of MCM1(1-100)

MCM1(1-100) was overexpressed using a pET3a-based expression plasmid containing the entire MCM1 coding region and a TAA stop codon inserted after MCM1 codon 100. BL21(DE3)pLyS cells containing the pET3a-MCM1(1-100) plasmid were induced at 27-29 °C with 0.2 to 0.4 mM IPTG and harvested after 10-15 hours incubation at the same temperature.

Cells from 6 l of culture overproducing MCM1(1-100) were resuspended in 150 ml of T300 (20 mM Tris-HCl (pH 8.0), 300 mM NaCl, 0.5 mM EDTA, 1 mM benzamidine, 10 mM 2-mercaptoethanol), lysed by freeze-thawing and homogenized using an Ultra-Turrax blender device to reduce the viscosity of the cell extract. A cleared extract was prepared by centrifuging the crude extract at 40,000 *g* and re-extraction of the pellet with an additional 100 ml of T300 buffer. The cleared extract was applied by continuous flow to a 2.5 cm internal diameter × 8 cm (40 ml) heparin-Sepharose column for 12-16 hours to ensure efficient binding of MCM1 to the resin. The column was then washed with 100 ml of T300, 150 ml of TU300 (3 M urea) (T300 with 3 M urea), 100 ml of T300 and eluted with a linear gradient of 50 ml of T300 + 50 ml of T1500. Flow from the heparin-Sepharose column was directed to a 10 ml Matrex Blue A (Amicon) column equilibrated in T300 when the single large peak was detected. The Matrex Blue A flow-through fractions were pooled, dialyzed against 1 mM Hepes (pH 7.5), 300 mM NaCl, 5 mM 2-mercaptoethanol and further purified over a 21.5 mm internal diameter × 150 mm TSK SP-5PW HPLC column at pH 7.5 (linear gradient from 25O mM to 1000 mM NaCl). The SP-5PW pool was then dialyzed against 1 mM Hepes (pH 7.5), 300 mM $NaNO_3$, 0.5 mM DTT and concen-

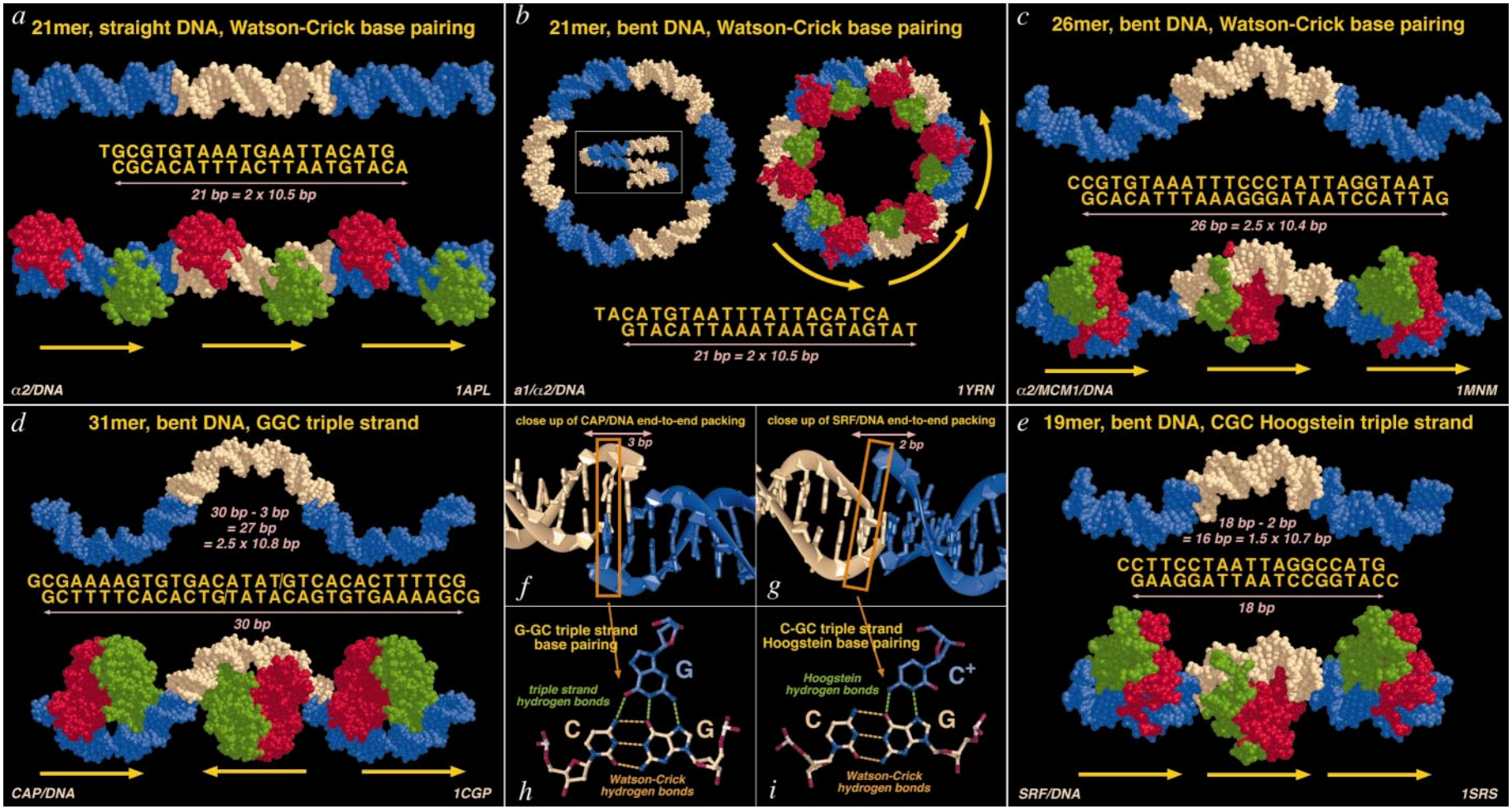
a
21mer, straight DNA, Watson-Crick base pairing
TGCGTGTAAATGAATTACATG
CGCACATTTACTTAATGTACA
21 bp = 2 x 10.5 bp
α2/DNA
1APL
b
21mer, bent DNA, Watson-Crick base pairing
TACATGTAATTTATTACATCA
GTACATTAAATAATGTAGTAT
21 bp = 2 x 10.5 bp
a1/α2/DNA
1YRN
c
26mer, bent DNA, Watson-Crick base pairing
CCGTGTAAATTTCCCTATTAGGTAAT
GCACATTTAAAGGGATAATCCATTAG
26 bp = 2.5 x 10.4 bp
α2/MCM1/DNA
1MNM
d
31mer, bent DNA, GGC triple strand
30 bp - 3 bp
= 27 bp
= 2.5 x 10.8 bp
GCGAAAAGTGTGACATATGTCACACTTTTCG
GCTTTTCACACTGTATACAGTGTGAAAAGCG
30 bp
CAP/DNA
1CGP
close up of CAP/DNA end-to-end packing
3 bp
f
close up of SRF/DNA end-to-end packing
2 bp
g
G-GC triple strand
base pairing
triple strand
hydrogen bonds
G
C
G
Watson-Crick
hydrogen bonds
h
C-GC triple strand
Hoogstein base pairing
Hoogstein
hydrogen bonds
C+
C
G
Watson-Crick
hydrogen bonds
i
e
19mer, bent DNA, CGC Hoogstein triple strand
18 bp - 2 bp
= 16 bp = 1.5 x 10.7 bp
CCTTCCTAATTAGGCCATG
GAAGGATTAATCCGGTACC
18 bp
SRF/DNA
1SRS

Figure 5

trated by ultrafiltration to 50-100 mg/ml. Yields varied between 100 and 250 mg from 6 l of cells. Protein concentrations were calculated using the theoretical extinction coefficient of 0.12 $(mg/ml)^{-1}$ cm^{-1} (Gill & von Hippel, 1989).

Overexpression and purification of α2(103-189)

The α2(103-189) gene fragment was subcloned into the pRET3a expression vector (S.T. and T.J.R., unpublished data), a pET3a derivative which expresses the ArgU gene to overcome potentially limiting amounts of $tRNA^{Arg}_{AGA,AGG}$ in *E. coli*. The α2(103-189) gene fragment was overexpressed in BL21(DE3)pLysS cells after two hours induction with 0.2 mM IPTG at 37 °C.

Cells overproducing α2(103-189) from 6 l of culture were resuspended in 150 ml of T100 buffer (20 mM Tris-HCl (pH 8.0) 0.5 mM EDTA, 1 mM benzamidine, 1 mM DTT, 100 mM NaCl), lysed by freeze-thawing and homogenized using an Ultra-Turrax blender device. A cleared extract was prepared by centrifuging the crude extract at 40,000 ***g*** and subsequent re-extraction of the pellet with 100 ml of T300 buffer. The cleared extract was applied in a single pass to a 2.5 cm internal diameter × 8 cm (40 ml) heparin-Sepharose column, and washed with 100 ml of T100, 100 ml of TU100 with 3 M urea, 100 ml of T100 and 100 ml of T200 buffer before elution with a linear gradient of 40 ml of T200 + 40 ml of T1500. After addition of ammonium sulfate to 1.5 M, the heparin-Sepharose pool was fractionated over a 21.5 mm internal diameter × 150 mm TSK Phenyl-5PW HPLC column at pH 8.0 (linear gradient from 1050 mM to 300 mM $(NH_4)_2SO_4$). The resulting phenyl pool was dialyzed against 1 mM Hepes (pH 7.5), 0.1 mM EDTA, 0.2 mM DTT, 100 mM NaCl prior to further purification over a 21.5 mm internal diameter × 150 mm TSK SP-5PW HPLC column at pH 7.5 (linear gradient from 370 mM to 460 mM NaCl). The purified α2(103-189) in the SP pool was dialyzed against 2 mM Hepes (pH 7.5), 0.1 mM EDTA, 0.5 mM DTT, 50 mM NaCl and concentrated against the same buffer by ultrafiltration to a final concentration of 10-30 mg/ml (yields of 15 to 40 mg from 6 l of cells).

Construction of plasmids containing multiple copies of STE2 UAS

The oligonucleotides STO47 (GGGTAAGTACAT GATATCATAAATTTCTAATT) and STO48 (ACCCAA TTAGGAAATTTACATGATATCATGTACTT) were purified by denaturing gel electrophoresis, kinased and annealed prior to self-ligation to create a ladder of fragments. Since STO47 and STO48 anneal to leave complementary but asymmetric sticky ends, the inserts in the ligated fragments should remain in the same orientation. The ligation mixture was treated with Klenow enzyme and fractionated over a 7.5% polyacrylamide gel. Nucleic acids from bands corresponding to 12 ×, 15× and 20× monomers were eluted and ligated into the *Sma*I site of pUC9. One of the clones obtained from ligat-

Figure 5. DNA packing in crystals of protein/DNA complexes. (a) Packing of an integral helical repeat (21 bp) of straight DNA using Watson-Crick base-pairing at the ends as represented by the α2/DNA complex (Wolberger *et al*., 1991). The top half shows the DNA (beige and blue) alone, while the bottom half shows the arrangement of the two protein monomers (red and green). Each complex is arranged in a direct tandem repeat, as signified by the yellow arrows. The sequence of the 21mer fragment is shown. PDB ID code is shown in the lower left. (b) Packing of an integral helical repeat (21 bp) of bent DNA using Watson-Crick base-pairing at the ends as represented by the **a**1/α2/DNA complex. The packing of the DNA alone is shown on the left, with a side view of the same in the insert to show that DNA forms a superhelix and not a closed circle. Each complex is arranged in a direct tandem repeat, as visible from the protein arrangement shown on the right (Li *et al*., 1995b). (c) Packing of a half-integral helical repeat (26 bp) of bent DNA using Watson-Crick base pairing at the ends as represented by the α2/MCM1/DNA complex (Tan & Richmond, 1998). The DNA packing on the top half shows a zig-zag arrangement of DNA monomers which form end-to-end packing *via* the complementary base overhang at either end. Direct tandem repeat of complexes is present here as well, as seen in the arrangement of the MCM1 polypeptide monomer on the DNA shown in the bottom half. The α2 polypeptides have been omitted for clarity. (d) Packing of half-integral helical repeat (27 bp) of bent DNA using G-GC triple stranded ends as represented by the CAP/DNA complex (Schultz *et al*., 1991). The two identical DNA strands contain a 13 and 18 base synthetic oligonucleotide, and the protein/DNA complex contains two nicks, signified by a slash in the diagram. Although both strands contain a nominal 31 base oligonucleotide (ignoring the nick), the DNA length is considered to be 30 bp, since the G overhang used for the triple strand pairing does not contribute to DNA length in the crystal. The 3 bp displacement of each DNA monomer with respect to its neighbor causes a loss of 3 bp from the nominal 30 bp, resulting in a 27 bp repeat or 2.5 turns of near *B*-form DNA arranged in a zig-zag array. The protein/DNA complexes are arranged in a head-to-head, tail-to-tail fashion, as can be seen from the arrangement of CAP monomers with respect to the DNA in the lower half. Both tandem repeats and inverted repeat arrangements are compatible with G-GC triple strand ends. (e) Packing of half-integral helical repeat (18 bp) of bent DNA using C-GC triple stranded ends as represented by the SRF/DNA complex (Pellegrini *et al*., 1995). Similar to the CAP/DNA complex, the DNA length is considered to be 18 bp, even though each strand contains 19 bases, since the C overhang used for the triple strand pairing does not contribute to DNA length in the crystal. The 2 bp displacement of each DNA monomer with respect to its neighbor causes a loss of 2 bp from the nominal 18 bp, resulting in a 16 bp repeat or 2.5 turns of near *B*-form DNA in a zig-zag array. The protein/DNA complexes are arranged in a tandem direct repeat, as can be seen from the arrangement of SRF polypeptide monomers in the lower half. (f) Close up of the end-to-end packing in CAP/DNA crystal. The displacement of each helix by 3 bp from its neighbor is shown. (g) Close up of the end-to-end packing in SRF/DNA crystal. The displacement of each helix by 2 bp from its neighbor is shown. (h) G-GC triple strand pairing of boxed region from (f). Watson-Crick hydrogen bonds are shown in orange while triple-strand hydrogen bonds are shown in green. (i) C^+-GC triple strand pairing of boxed region from (g). Note that the triple strand cytosine base must be protonated for these hydrogen bonds to occur. Same color code for hydrogen bonds as in (h).

ing the 12× monomer band, pST8, contained four copies of the 35mer monomer in tandem head-to-tail repeats. Other clones contained three or fewer copies of tandem repeats.

The 105 bp *Eco*RV fragment containing three tandem head-to-tail repeats of the 35mer was isolated by partial digestion of pST8 and ligated into the *Eco*RV site of a pUC9 derivative containing a *Sal*I-*Eco*RV-*Xho*I polylinker. One of the resulting clones contained, fortuitously, two tandem copies of the 105 bp *Eco*RV fragment equivalent to six copies of the 35mer. This insert in this plasmid, pST11, could be ''doubled'' by using the compatible *Sal*I and *Xho*I ends: the *Eco*RI-*Xho*I insert fragment containing 6×35mer was ligated with the *Eco*RI-*Sal*I vector fragment containing 6×35mer to produce pST12 containing 12 copies of the 35mer. This doubling process was repeated twice to yield pST13 with 24 copies of the 35mer, and pST17 with 48 copies of the 35mer. pST13, but not pST17, was apparently stable in the HB101, and stability of pST17 was improved by maintaining in the SURE strain (Stratagene). Attempts to construct plasmids containing 72 or 96 copies of the 35mer yielded plasmids with unstable inserts in both HB101 and SURE strains.

Isolation and purification of STE2 35mer

The STE2 35mer was isolated from either HB101/pST13 or SURE/pST17 cells using the same procedure, with similar yields (although pST17 contains more insert copies than pST13, the yield of plasmid from SURE cells was lower than HB101 in the single SURE/pST17 preparation performed). A modified version of published procedures by Richmond *et al*. (1988), incorporating the use of a higher concentration of acetate using the alkaline lysis procedure (Feliciello & Chinali, 1993), was used to isolate plasmid DNA from 12 l of culture.

Purified plasmid was digested overnight with *Eco*RV restriction endonuclease to release the 35mer monomer. Fractionation of the vector from the insert in the digestion mix was achieved by precipitation with 9% PEG6000 in the presence of the 0.5 M NaCl on ice, followed by centrifugation at 9500 *g* at 4 °C. The 35mer in the supernatant was precipitated by addition of 2.5 volumes of absolute ethanol and incubation at 20 °C for 15 minutes. The precipitate was collected by centrifugation at 9000 rpm and 20 °C, and resuspended in 20 mM Tris-HCl (pH 7.6) 0.1 mM EDTA, 300 mM NaCl. The 35mer was purified by anion-exchange chromatography over a 21.5 mm internal diameter × 150 mm TSK DEAE-5PW column at pH 7.6 using a linear gradient from 390 mM to 450 mM NaCl.

The pooled fractions were precipitated by addition of 2.5 volumes of absolute ethanol and $MgCl_2$ to 10 mM and incubation at 20 °C for 15 minutes. The precipitate collected by centrifugation at 9500 *g* for 20 minutes was resuspended in 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA to a final concentration of 40 to 50 mg/ml. Preparation of plasmid DNA from 12 l requires one to two days, while purification of STE2 35mer from the plasmid requires two to three additional days.

Purification of oligonucleotides for crystallization

Oligonucleotides synthesized on the 1 μmol scale (Applied Biosystems 380B) were purified by reversed-phase HPLC, including detritylation during the single chromatography run. The oligonucleotides were synthesized with the 5′ trityl protecting group left on. After heat deprotection, 100 μl of 1 M Tris base (pH not adjusted) was added to the sample, dried in a Speed Vac and resuspended in 800 μl of 0.1 M TEA-Ac (pH 6.7). The sample was loaded onto a Nucleosil 300-5 C8 reversed-phase HPLC column equilibrated in 15% CH_3CN in 0.1 M TEA-Ac (pH 6.7). After the failed synthesis products had eluted, the column was washed with 0.1 M TEA-Ac (pH 6.7). The bound oligonucleotides were detritylated while bound to the column by two injections of 5 ml of 0.5% (v/v) TFA. The detritylated oligonucleotides were then eluted off the column using a gradient of 10% to 16% CH_3CN for oligonucleotides between 16 and 25 bases long (gradients of 8% to 14% or 11.2% to 17.2% CH_3CN was used for 10 to 15mers or 25 to 28mers, respectively). The main eluted peak was pooled, dried in a Speed Vac, resuspended in 300 μl of 1 M NaCl and precipitated by addition of 1 ml of absolute ethanol and centrifugation in a microfuge for ten minutes. The precipitated oligonucleotide was then resuspended in 500 μl of water and lyophilized to dryness before being dissolved in 150 μl of 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA. Oligonucleotide pairs were annealed together in equimolar ratios by heating to 90 °C for three minutes and gradual cooling to room temperature over longer than one hour. $MgCl_2$ was added to final concentration of 10 mM for annealing of symmetric oligonucleotides to increase the efficiency of annealing.

The same procedure was used for oligonucleotides containing iodinated nucleotides (Glen Research; Sterling, VA) except that deprotection was performed for 24 hours at room temperature and samples were kept in the dark wherever possible.

Purification and crystallization of MCM1(1-100)/DNA complexes

MCM1(1-100) was mixed in equimolar ratio with DNA and used directly for crystallization trials. Alternatively, MCM1(1-100) polypeptide was combined with DNA in 10 mM Hepes (pH 7.5), 5 mM $MgCl_2$, 60 mM KCl, 5% (v/v) glycerol in a final volume of 300 μl and the sample purified by preparative gel electrophoresis (model 491 Prep Cell, Bio-Rad) using a 6% polyacrylamide (40:1 (w/w), acrylamide/bis-acrylamide) gel. Some 5 mg of complex can be purified in a single electrophoretic run. The pooled fractions of 10 to 15 ml eluted with 20 mM $NaNO_3$, 0.5 mM DTT were concentrated by ultrafiltration against the same buffer and used for crystallization trials.

Hanging drop trials were used for most crystallization attempts. Crystals of MCM1(1-100)/STO138/139 were obtained by equilibrating the protein/DNA complex in 10 mM Bis-Tris (pH 6.5), 100 mM $NaNO_3$, 5 mM $Mg(NO_3)_2$, 5 mM DTT and 3-5% PEG4000 against a reservoir of 5 mM sodium acetate (pH 5.2) to pH 5.5, 200 mM $NaNO_3$, 10 mM $Mg(NO_3)_2$, 10 mM DTT and 6-10% PEG4000 at 4 to 8 °C.

Purification and crystallization of α2(103-189)/MCM1(1-100)/DNA complexes

Ternary complexes of α2(103-189), MCM1(1-100) and DNA were formed by mixing the individual components in equimolar ratios in 10 mM Bis-Tris (pH 7.0), 100 mM NaCl, 5 mM DTT, 5% glycerol. Purification of the ternary complex was achieved by gel filtration over 10 mm internal diameter × 300 mm Superose 12 or Superdex

200 HR columns eluted with 10 mM Bis-Tris (pH 7.0), 100 mM NaCl, 0.1 mM EDTA, 0.5 mM DTT at 0.4 ml/minute. The pooled fractions were concentrated by ultrafiltration against the elution buffer to a final concentration of 10 mg/ml and crystallized in hanging drop against 50 mM sodium acetate (pH 5.5), 100 mM ammonium acetate, 10 mM $CaCl_2$, 7-10% PEG6000 (complex mixed 1:1 (v/v) with reservoir buffer).

References

Acton, T. B., Zhong, H. & Vershon, A. K. (1997). DNA-binding specificity of Mcm1: operator mutations that alter DNA-bending and transcriptional activities by a MADS box protein. *Mol. Cell Biol.* **17**, 1881-1889.

Anderson, J., Ptashne, M. & Harrison, S. C. (1984). Cocrystals of the DNA-binding domain of phage 434 repressor and a synthetic phage 434 operator. *Proc. Natl Acad. Sci. USA,* **81**, 1307-1311.

Bruhn, L., Hwang-Shum, J. J. & Sprague, G. F., Jr. (1992). The N-terminal 96 residues of MCM1, a regulator of cell type-specific genes in *Saccharomyces cerevisiae*, are sufficient for DNA binding, transcription activation, and interaction with alpha 1. *Mol. Cell Biol.* **12**, 3563-3572.

Christ, C. & Tye, B. K. (1991). Functional domains of the yeast transcription/replication factor MCM1. *Genes Dev.* **5**, 751-763.

Feliciello, I. & Chinali, O. (1993). A modified alkaline lysis method for the preparation of highly purified plasmid DNA from *Escherichia coli*. *Anal. Biochem.* **212**, 394-401.

Flaus, A., Luger, K., Tan, S. & Richmond, T. J. (1996). Mapping nucleosome position at single base-pair resolution by wing site-directed hydroxyl radicals. *Proc. Natl Acad. Sci. USA,* **93**, 1370-1375.

Fountoulakis, M., Takacs-di Lorenzo, B., Juranville, J. F. & Manneberg, M. (1993). Purification of interferon gamma-interferon gamma receptor complexes by preparative electrophoresis on native gels. *Anal. Biochem.* **208**, 270-276.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.

Gupta, S. C., Weith, H. L. & Somerville, R. L. (1983). Biological limitations on the length of highly repetitive DNA sequences that may be stably maintained within plasmid replicons in *Escherichia coli*. *Bio/Technol.* **1**, 602-609.

Harp, J. M., Palmer, E. L., York, M. H., Gewiess, A., Davis, M. & Bunick, G. J. (1995). Preparative separation of nucleosome core particles containing defined-sequence DNA in multiple translational phases. *Electrophoresis,* **16**, 1861-1864.

Jin, Y., Mead, J., Li, T., Wolberger, C. & Vershon, A. K. (1995). Altered DNA recognition and bending by insertions in the alpha 2 tail of the yeast a1/alpha 2 homeodomain heterodimer. *Science,* **270**, 290-293.

Jordan, S. R., Whitcombe, T. V., Berg, J. M. & Pabo, C. O. (1985). Systematic variation in DNA length yields highly ordered repressor-operator cocrystals. *Science,* **230**, 1383-1385.

Li, T., Stark, M., Johnson, A. D. & Wolberger, C. (1995a). Crystallization and preliminary X-ray diffraction studies of an a1/alpha 2/DNA ternary complex. *Proteins: Struct. Funct. Genet.* **21**, 161-164.

Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. (1995b). Crystal structure of the MATa1/MAT alpha 2 homeodomain heterodimer bound to DNA. *Science,* **270**, 262-269.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997a). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* **389**, 251-260.

Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M. & Richmond, T. J. (1997b). Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* **272**, 301-311.

Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R. & Sigler, P. B. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature,* **352**, 497-505.

Palmer, E. L., Gewiess, A., Harp, J. M., York, M. H. & Bunick, G. J. (1995). Large-scale production of palindrome DNA fragments. *Anal. Biochem.* **231**, 109-114.

Pellegrini, L., Tan, S. & Richmond, T. J. (1995). Structure of serum response factor core bound to DNA. *Nature,* **376**, 490-498.

Primig, M., Winkler, H. & Ammerer, G. (1991). The DNA binding and oligomerization domain of MCM1 is sufficient for its interaction with other regulatory proteins. *EMBO. J.* **10**, 4209-4218.

Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. (1996). Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell,* **87**, 1295-1306.

Richmond, T. J., Searles, M. A. & Simpson, R. T. (1988). Crystals of a nucleosome core particle containing defined sequence DNA. *J. Mol. Biol.* **199**, 161-170.

Schultz, S. C., Shields, G. C. & Steitz, T. A. (1990). Crystallization of *Escherichia coli* catabolite gene activator protein with its DNA binding site. The use of modular DNA. *J. Mol. Biol.* **213**, 159-166.

Smith, D. L., Desai, A. B. & Johnson, A. D. (1995). DNA bending by the a1 and alpha 2 homeodomain proteins from yeast. *Nucl. Acids Res.* **23**, 1239-1243.

Tan, S. & Richmond, T. J. (1998). Crystal structure of the yeast MATalpha2/MCM1/DNA ternary complex. *Nature,* **391**, 660-666.

Vershon, A. K. & Johnson, A. D. (1993). A short, disordered protein region mediates interactions between the homeodomain of the yeast alpha 2 protein and the MCM1 protein. *Cell,* **72**, 105-112.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. (1991). Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell,* **67**, 517-528.

Yang, W. & Steitz, T. A. (1995). Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. *Cell,* **82**, 193-207.

Edited by K. Nagai

(Received 23 November 1999; received in revised form 8 February 2000; accepted 15 February 2000)