

Structures of the hyperthermophilic chromosomal  
protein Sac7d in complex with DNA decamersTzu-Ping Ko,<sup>a</sup> Hsing-Mao Chu,<sup>b</sup>  
Chin-Yu Chen,<sup>c</sup> Chia-Cheng  
Chou<sup>d</sup> and Andrew H.-J.  
Wang<sup>a,b,d,\*</sup><sup>a</sup>Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan, <sup>b</sup>Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan, <sup>c</sup>Department of Chemistry, National Taiwan University, Taipei 106, Taiwan, and <sup>d</sup>Core Facility for Protein X-ray Crystallography, Academia Sinica, Taipei 115, TaiwanCorrespondence e-mail:  
ahjwang@gate.sinica.edu.tw

The protein Sac7d belongs to a class of small chromosomal proteins from the hyperthermophilic archaeon *Sulfolobus acidocaldarius*. Two new crystal forms of Sac7d in complex with the DNA decamers CCTATATAGG and CCTACGTAGG were obtained and their structures were determined by molecular replacement. The refined models yielded  $R/R_{\text{free}}$  values of 0.221/0.257 and 0.248/0.290 at 1.9 and 2.2 Å resolution, respectively. The protein structures are similar to the previously determined structure of Sac7d–GCGATCGC (PDB code 1azp), but the DNA molecules are more bent overall, by 14–20°. The relative positions of the Sac7d protein and the bound DNA also differ by rotations of 6–10° and translations of 1.0–2.4 Å. In addition to the water molecules in the central cavity, three additional conserved water molecules are found that mediate the protein–DNA interactions. The decamer DNA fragments form virtual double helices in the crystal, with a unit length of eight base pairs. The molecular packing of the new crystal forms differs from that of 1azp. The terminal nucleotides are opened up and form triple base pairs with other DNA molecules. Through lattice contacts, the Sac7d molecule also makes additional interactions with DNA, whereas only limited protein–protein interactions are seen.

Received 25 March 2004  
Accepted 18 May 2004

## PDB References:

Sac7d–CCTATATAGG,  
1wd0, r1wd0sf;  
Sac7d–CCTACGTAGG,  
1wd1, r1wd1sf.

## 1. Introduction

The archaeobacterium *Sulfolobous acidocaldarius*, which lives in hot springs, protects its DNA from denaturation by binding to heat-stable proteins (Reddy & Suryanarayana, 1988). Sac7d belongs to a class of small chromosomal DNA-binding proteins isolated from *S. acidocaldarius* (Choli *et al.*, 1988). It comprises 66 amino-acid residues and has a total weight of 7.6 kDa. The solution structure of Sac7d was first determined by NMR spectroscopy (Edmondson *et al.*, 1995), which showed a compact folding with a C-terminal  $\alpha$ -helix packed against an N-terminal  $\beta$ -sheet consisting of five antiparallel strands. X-ray diffraction analysis of the crystal structure of Sac7d in complex with DNA further showed that the protein binds to DNA through the minor groove, while intercalation of the side chains of Val26 and Met29 into adjacent base pairs produces a sharp kink of about 72° in the DNA double helix (Robinson *et al.*, 1998). Sac7d increases the melting temperature of DNA by nearly 40 K *in vitro*, with little preference for a specific nucleotide sequence. Presumably, this will stabilize duplex DNA and help its packaging in the bacterial cell. The structure of Sso7d, a homologous protein from *S. solfataricus*, was also determined in complex with DNA and showed similar effects on bending DNA, despite the use of different crystal forms (Gao *et al.*, 1998).

In eukaryotic cells, the chromosomal DNA molecules are bound to histones and are organized into chromatin. The

**Table 1**

X-ray data and model statistics for Sac7d–DNA decamer crystals.

Values in parentheses are for the highest resolution shells.

Data collection	CCTATATAGG	CCTACGTAGG
DNA sequence	CCTATATAGG	CCTACGTAGG
Unit-cell parameters		
<i>a</i> (Å)	38.97	38.63
<i>b</i> (Å)	50.91	46.84
<i>c</i> (Å)	51.99	53.33
Resolution range (Å)	30–1.9 (1.97–1.90)	30–2.2 (2.28–2.20)
No. observations	55758 (3414)	24387 (1638)
Unique reflections	8591 (827)	5166 (453)
Completeness (%)	99.7 (97.5)	98.0 (88.5)
Average <i>I</i> / $\sigma$ ( <i>I</i> )	43.3 (2.3)	24.9 (2.0)
<i>R</i> <sub>merge</sub> (%)	4.5 (49.0)	6.5 (54.1)
Refinement		
Total reflections used [ <i>F</i> > 0 $\sigma$ ( <i>F</i> )]	8203 (691)	4798 (386)
<i>R</i> for 95% working data set	0.221 (0.324)	0.248 (0.389)
<i>R</i> <sub>free</sub> for 5% test data set	0.257 (0.307)	0.290 (0.476)
R.m.s.d. from ideal bond lengths (Å)	0.019	0.020
R.m.s.d. from ideal bond angles (°)	1.93	2.17
Ramachandran plot: No. residues		
In most favoured regions (%)	89.7	85.7
In additional allowed regions (%)	8.6	14.3
Average <i>B</i> values/No. atoms		
Protein (Å <sup>2</sup> )	38.8/533	30.2/515
DNA (Å <sup>2</sup> )	36.3/404	32.7/372
Water (Å <sup>2</sup> )	46.6/95	44.5/89

mechanism of DNA packaging in the archaeobacterial cell, however, remains to be further investigated (White & Bell, 2002). It was observed in the crystal structures of Sac7d–DNA complexes (Robinson *et al.*, 1998; Gao *et al.*, 1998; Su *et al.*, 2000) that the DNA fragments stacked with symmetry-related neighbours *via* the 3'- and 5'-ends and formed virtual double helices that spanned the entire crystals. The protein molecules were bound in an alternating manner along the DNA and kinked the DNA back and forth. No extensive interaction was observed between the Sac7d molecules. In this study, we tried to grow crystals using DNA molecules of ten and 12 base pairs, which are longer than those previously used. The protein:DNA ratio was also increased in order to offer the possibility for two Sac7d molecules to bind to the same DNA. However, it turned out that the crystals still contained 1:1 protein–DNA complexes, but some interesting new features of DNA kinks and crystal contacts were seen.

## 2. Materials and methods

### 2.1. Crystallization and data collection

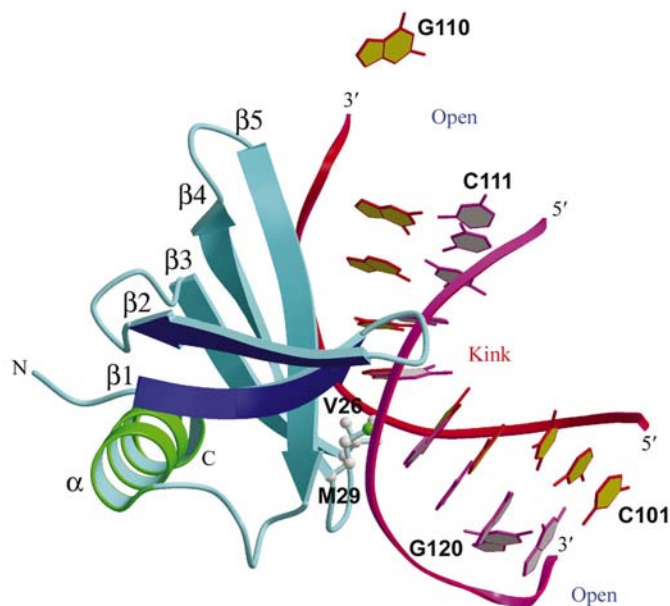
The Sac7d protein was purified as described previously (McAfee *et al.*, 1995). The procedure included heating, acid treatment and ion-exchange chromatography using a HiTrap CM-FF column (Amersham Pharmacia). After dialysis against distilled water, the protein was lyophilized and stored at 253 K. Prior to crystallization, the protein and DNA were dissolved in distilled water to final concentrations of 10 and 5 mM, respectively, and mixed in a 1:1 ratio. Hanging drops were set up by mixing 1  $\mu$ l of the Sac7d–DNA solutions with 1  $\mu$ l reservoir solution and were equilibrated against reservoir solution at room temperature. Crystals of Sac7d–CCTATATAGG were obtained using reservoir solutions containing

2 mM Tris–HCl pH 6.5 and 8% polyethylene glycol (PEG 400) and crystals of Sac7d–CCTACGTAGG were obtained using 50 mM KH<sub>2</sub>PO<sub>4</sub> pH 4 and 10% PEG 8000.

Data collection was carried out at 123 K using a MicroMax 002 X-ray generator and an R-Axis IV<sup>++</sup> (Molecular Structure Corporation) detector. Prior to flash-freezing, crystals of Sac7d–CCTATATAGG and Sac7d–CCTACGTAGG were soaked in 30% glycerol and 25% PEG 200, respectively, as cryoprotectants. The diffraction images were processed using the *HKL* package (Otwinowski & Minor, 1997). Both crystal forms belong to the orthorhombic space group *P*2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, with similar unit-cell parameters, and both contain one Sac7d–DNA complex in an asymmetric unit. However, the two crystal forms are not isomorphous. Statistics for the two data sets are shown in Table 1.

### 2.2. Structure determination and refinement

The structure of the Sac7d–CCTATATAGG crystal was solved by molecular replacement using the program *CNS* (Brünger *et al.*, 1998). A data set collected earlier to 2.2 Å resolution was employed and the Sac7d–GCGATCGC complex (Robinson *et al.*, 1998; PDB code 1azp) was used as a search model. The initial *R* value was 0.47 after rigid-body refinement; preliminary cycles of energy minimization yielded an *R* value of 0.38. When a 2*F*<sub>o</sub> – *F*<sub>c</sub> electron-density map was calculated, additional nucleotides were observed and a DNA model with the correct sequence was constructed. The program *O* (Jones *et al.*, 1991) was employed for viewing and manipulation of the model. Water molecules were added according to a density level of 1.0 $\sigma$  in the 2*F*<sub>o</sub> – *F*<sub>c</sub> map. The



**Figure 1** Ribbon representation of the Sac7d–CCTATATAGG complex. The structure is shown with five  $\beta$ -strands and one  $\alpha$ -helix for the protein and with eight base pairs for the DNA. The four terminal bases, however, are open. The DNA is sharply kinked at the central TA step by the residues Val26 and Met29 in the  $\beta$ 3– $\beta$ 4 hairpin. This figure and all other figures were prepared using *MolScript* (Kraulis, 1991) and *Raster3D* (Merritt & Murphy, 1994).

N- and C-terminal residues Met1 and Glu64–Lys66 of the protein and the 3'-G of both DNA strands lacked clear density and were omitted. Subsequent refinement yielded  $R$  and  $R_{\text{free}}$  values of 0.244 and 0.293 with the incorporation of 74 water molecules. Using the new data set, the  $R$  value at 2.2 Å resolution was 0.41 and decreased to 0.29 after rigid-body refinement. The resolution was then increased to 1.9 Å in subsequent refinements and map calculations. Later, electron densities were observed for the protein termini and the 3' nucleotides and these were added to the model.

The crystals of Sac7d–CCTATATAGG and Sac7d–CCTACGTAGG are not isomorphous, although their unit-cell parameters are similar and they both belong to the same space group. Consequently, the structure of Sac7d–CCTACGTAGG also had to be determined by molecular replacement, in which the refined Sac7d–CCTATATAGG structure was used as a search model. The  $R$  value for all data at 2.2 Å resolution was initially 0.50 and fell to 0.44 after rigid-body refinement. The N- and C-terminal residues Met1, Lys65 and Lys66 of the protein were in severe contact with symmetry-related molecules, as were the nucleotides 5'-C and 3'-G of both DNA strands. After their removal, the  $R$  value was 0.42 and an initial  $2F_o - F_c$  electron-density map was calculated. The two central TA base pairs were substituted with CG as clearly seen in the  $2F_o - F_c$  map. The N-terminal residue of Met1 was deleted owing to lack of corresponding density. On the other hand, the 3'-G of the first DNA strand and the 5'-C of the second, as well as the C-terminal Lys65 of the protein, were added back into the model as implicated by the electron-density map.

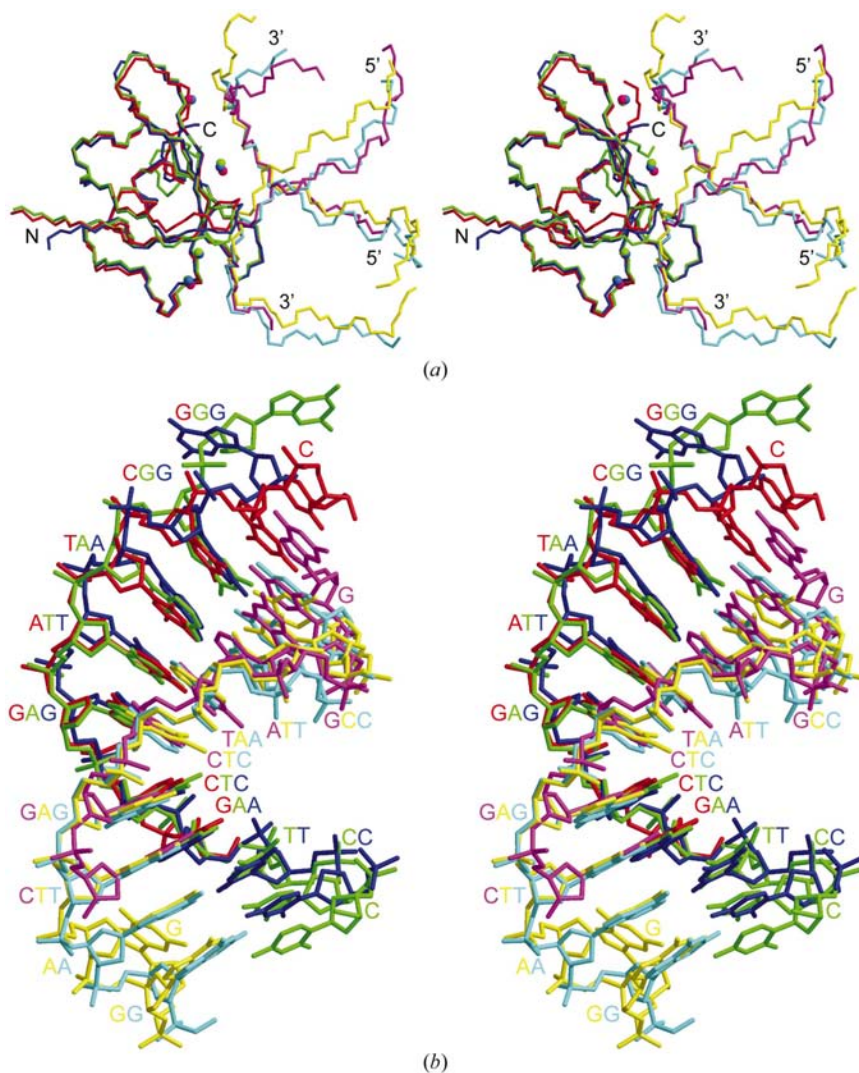
### 3. Results and discussion

#### 3.1. Overall structure

The final model of the Sac7d–CCTATATAGG complex (subsequently referred to as S7d-D1) contains the entire protein and DNA molecules, numbered Met1–Lys66 and C101–G120, plus 95 water molecules. A ribbon representation is shown in Fig. 1. The  $R$  and  $R_{\text{free}}$  values are 0.221 and 0.257 for all data at 1.9 Å resolution. At the C-terminus, 46 atoms in residues Arg60–Lys66 have higher temperature factors than  $70 \text{ \AA}^2$ , indicating some disorder. Only one residue, Lys65, has dihedral angles in the generously allowed regions defined by *PROCHECK* (Collaborative Computational Project, Number 4, 1994). The model of the Sac7d–CCTACGTAGG complex (subsequently referred to as S7d-D2) contains protein and

DNA residues Val2–Lys65 and C102–G119, plus 89 water molecules. The 5'- and 3'-phosphate groups of C102 and G119, respectively, are also included in the model.  $R$  and  $R_{\text{free}}$  are 0.248 and 0.290 at 2.2 Å resolution. Additional statistics are shown in Table 1.

The protein models of both crystals in this study superimpose well with that of the Sac7d–GCGATCGC complex (Robinson *et al.*, 1998; PDB code 1azp). Excluding the terminal residues Met1 and Lys66, the root-mean-square deviations (r.m.s.d.) of protein coordinates of S7d-D1 and S7d-D2 models from those of the 1azp model are 1.939 and 1.803 Å, respectively, for 515 atoms. Between S7d-D1 and S7d-D2 the r.m.s.d. is 1.720 Å. If side-chain atoms beyond CB are not included, the r.m.s.d.s are 1.088, 0.932 and 0.967 Å, respectively, for 315 atoms. The largest differences occur in the C-terminal regions, as shown in Fig. 2(a), with maximal



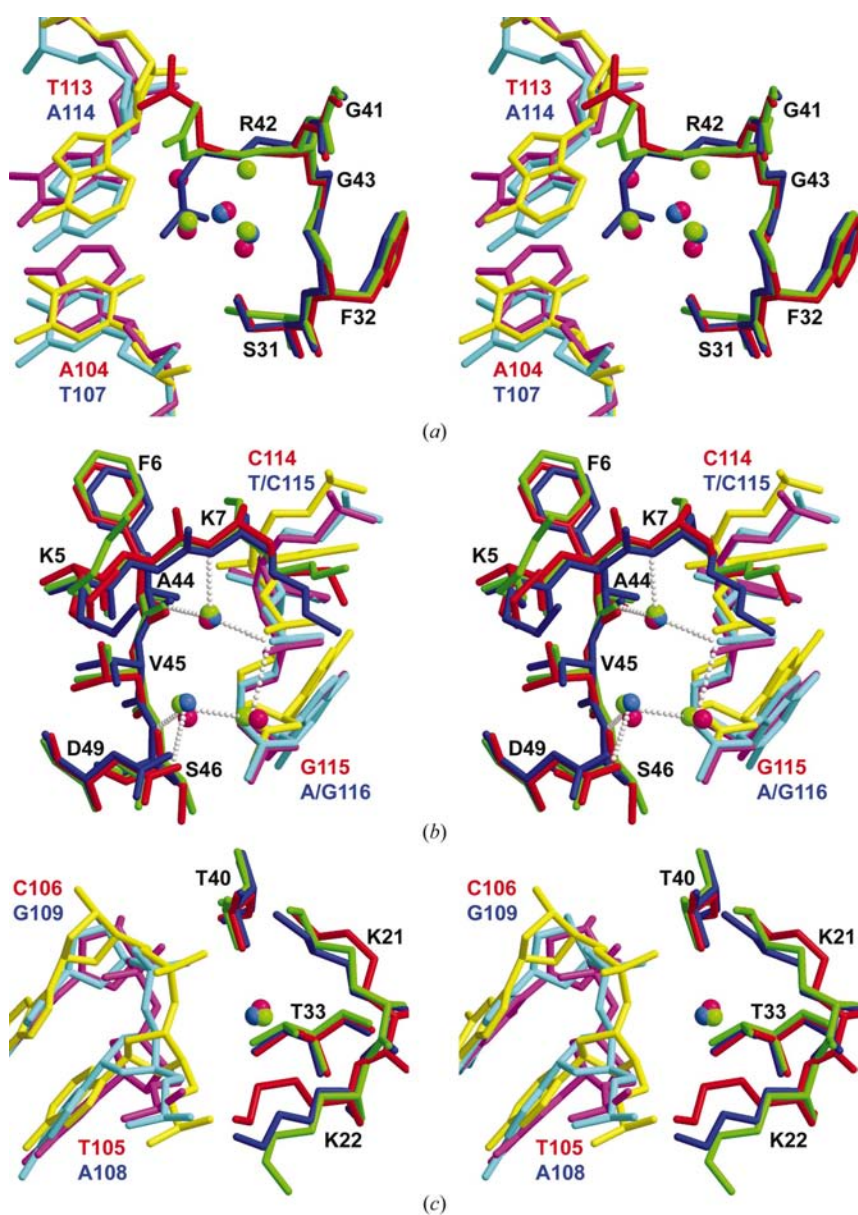
**Figure 2**

Comparison of the Sac7d–DNA structures in three crystal forms. In (a) the models are superimposed based on the protein moieties, with the polypeptide tracings coloured red, green and blue for the crystals of 1azp, S7d-D1 and S7d-D2, respectively. The DNA-backbone tracings are shown in magenta, yellow and cyan, respectively, and the four conserved water molecules are in similar colours. In (b) the DNA models alone are superimposed. The strands are coloured red/magenta, green/yellow and blue/cyan, respectively, for the three different crystals.



deviations of 3.8, 3.2 and 7.9 Å between the C $\alpha$  atoms of Glu64, Lys65 and Lys66, respectively. In the crystals of 1azp and S7d-D2, the C-termini are in contact with symmetry-related molecules, but in the S7d-D1 crystal the C-terminus protrudes into the bulk solvent. The N-termini are also exposed to solvent in all three crystals. However, the N-terminus has a similar disposition in the 1azp and S7d-D1 crystals, with a deviation of 0.8 Å between the C $\alpha$  atoms of Met1. In 1azp and S7d-D2 the C $\alpha$  deviation of Val2 is 2.9 Å. Two other regions with large C $\alpha$  deviations of 2.7 and 2.4 Å, respectively, are located in the loops Lys9–Glu12 and Asn37–Gly38. Again, these are involved in crystal contacts.

On the other hand, much larger differences are seen in the DNA molecules of the three Sac7d–DNA complexes. Unlike the octameric DNA of the 1azp crystal, in which all of the nucleotides make eight well defined base pairs, both decamer DNA molecules of the S7d-D1 and S7d-D2 crystals have their 5'- and 3'-ends opened up. The nucleosides of C101 and G120 in S7d-D2 were not visible in the electron-density maps. Nevertheless, the nucleotides C102–G109 and C112–G119 make eight regular base pairs. A104–G109 and C111–T117 of both decamers have similar conformations to G101–C106 and G110–C116 of the octamer. Excluding the base moieties, 141 atoms in the sugar-phosphate backbones of S7d-D1 and S7d-



**Figure 3**  
The conserved water molecules in the Sac7d–DNA complex structures. The models were overlapped and coloured as in Fig. 2(a). Protein and DNA models are shown as sticks and water molecules as spheres. In (a) the region near the centre of the protein–DNA interface, where several waters are found in a cavity, is shown. (b) and (c) show the regions where the other three conserved water molecules are located. In (b) potential hydrogen bonds are also shown, along with another water molecule found in the 1azp and S7d-D1 crystals that mediates the interactions between the DNA backbone and the protein.

D2 can be superimposed on those of 1azp, with r.m.s.d.s of 1.213 and 1.131 Å, respectively. As shown in Fig. 2(b), the nucleotides C101–T103 and A118–G120 in the decamers do not have equivalents in the 1azp structure. Between the S7d-D1 and S7d-D2 models, the 193 backbone atoms of nucleotides C102–G109 and C111–G119 deviate by 1.298 Å. The C1' atoms of G110 in S7d-D1 and S7d-D2, which do not form base pairs with C111, deviate by 3.6 Å and are 6.2 and 3.8 Å from those of G107 in 1azp. The guanine bases have disparate orientations, with a maximal deviation of 11.8 Å.

The DNA double helix of 1azp is kinked by 61.3° in the step C102–G103. (For simplicity, only the nucleotides in the first strand of DNA are mentioned.) Similar kinks of 48.6 and 62.2° occur in the S7d-D1 and S7d-D2 structures at T105–A106 and C105–G106, respectively. The single-step kink of T–A in S7d-D1 is significantly smaller than that of C–G in the other two structures, probably because of different stacking interactions between the base pairs. This is also expressed in the smaller values of the buckle angles of the base pairs T105/A116 and A106/T115. On the other hand, the overall bending angles of the DNA double helices in 1azp, S7d-D1 and S7d-D2 are 80, 100 and 94°, respectively, with S7d-D1 being the most bent. It appears that the perturbation in DNA by Sac7d binding is dissipated over a number of nucleotides in both directions. The longer AT stretch in D1 may account for the spreading of the kinks over several base pairs.

### 3.2. Protein–DNA interactions and the water models

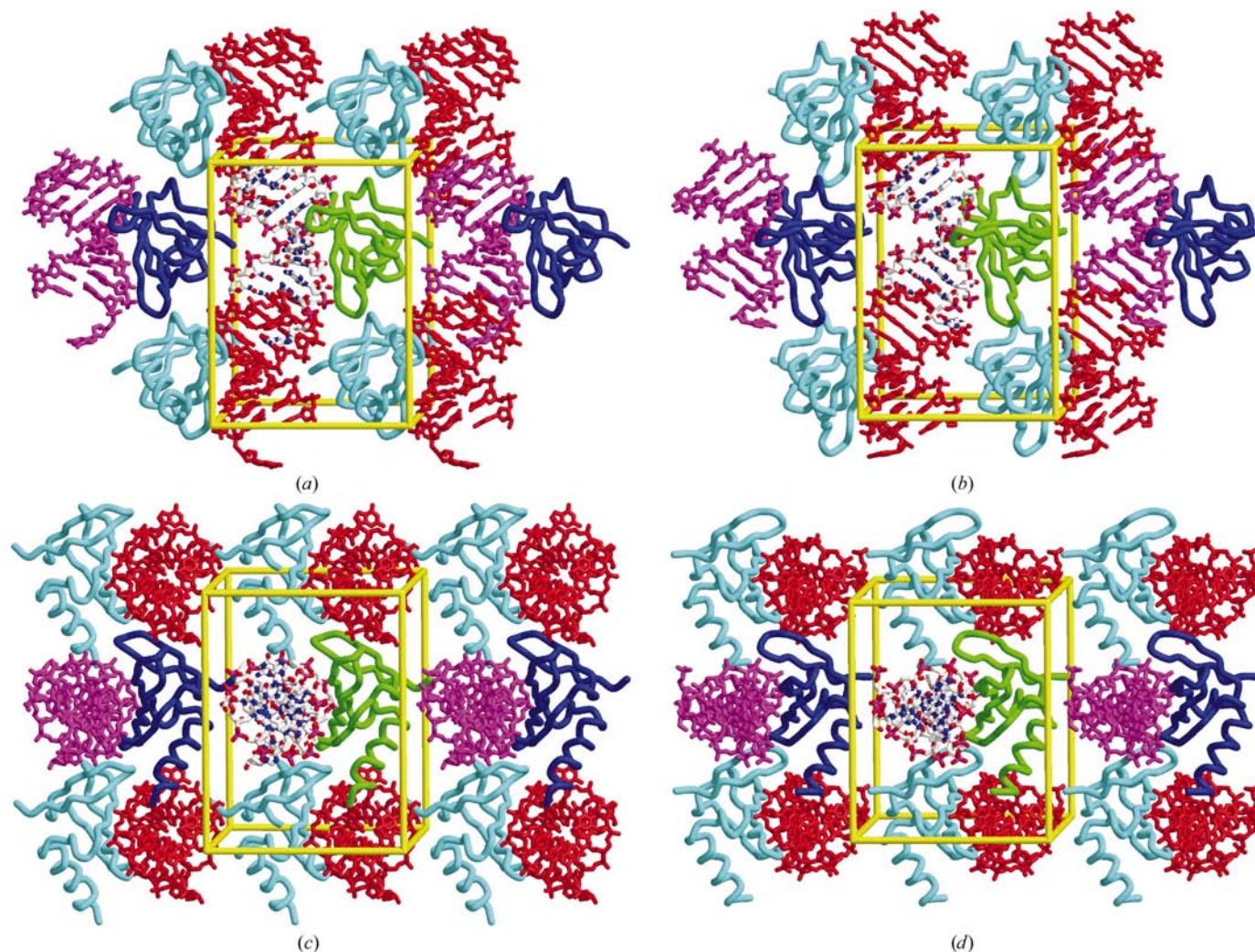
Most interactions between the Sac7d molecules and DNA octomers observed in previous crystal structures (Robinson *et al.*, 1998; Gao *et al.*, 1998; Su *et al.*, 2000) are also seen in both crystals with decameric DNA studied here. These include (i) salt bridges



between the DNA and the lysine residues 7, 9, 21, 22, 28, 39, and possibly 48 and 65, (ii) intercalation of the Val26 and Met29 side chains into the base pairs and (iii) hydrogen bonds between Trp24, Ser31, Thr33, Arg42, Ser46 and the DNA. Tyr8 and Ala44 are also involved. In Fig. 2(a), where the protein models are superimposed, more deviations are seen between the DNA models. In Fig. 2(b), however, the backbones of DNA also superimpose quite well, as do the equivalent bases in the central region. The difference in the disposition of the DNA molecule with respect to the protein is largest between 1azp and S7d-D1, which corresponds to a rotation angle of  $9.6^\circ$  plus a translation distance of 2.4 Å. The difference between 1azp and S7d-D2 corresponds to a rotation of  $5.8^\circ$  and a translation of 1.8 Å. Between S7d-D1 and S7d-D2 the angle and distance are  $6.4^\circ$  and 1.0 Å. Although Sac7d binds DNA without sequence specificity, the protein appears to interact differently with TA and GC steps of base pairs. As shown above, the matched backbones of 1azp and S7d-D2 are

more similar to each other than Sd7-D1. In particular, there is about a  $13^\circ$  difference in the kink angle.

There are four conserved water molecules in 1azp, S7d-D1 and S7d-D2, as shown in Fig. 2(a), where the distances between equivalents are all less than 1.0 Å. If the models are superimposed using the bound DNA molecules as in Fig. 2(b), no such equivalent water molecule can be found. All four water molecules are located at the protein–DNA interface and have identical interactions with protein and DNA. However, only one of these is located in the region of four interface waters reported as ‘always present’ in previous structures of Sac7d–DNA complexes (Su *et al.*, 2000). In the S7d-D1 crystal this region contains three water molecules, with distances of 1.6, 1.9 and 0.99 Å from the nearest waters in 1azp. In S7d-D2, only two water molecules, with distances of 0.25 and 0.71 Å, can be found. As shown in Fig. 3(a), one of the water positions in 1azp is too close to the N3 atom of A114 in S7d-D1 and S7d-D2, with distances of 1.9 and 2.4 Å, respectively. There-



**Figure 4**  
Packing diagrams of the Sac7d–DNA decamer complex crystals. Two orthogonal views are shown for both S7d-D1 (a and c) and S7d-D2 (b and d) crystals. In (a) and (b) the view is approximately parallel to the crystallographic *b* axis, with the *a* axis lying horizontal. In (c) and (d) the view is parallel to the *c* axis, with the *a* axis horizontal. An asymmetric unit includes a Sac7d molecule shown in green and a DNA double helix shown as stick models. Other symmetry-related molecules are shown with the protein and DNA coloured in cyan/blue and red/magenta, respectively. For clarity, only a layer of the molecules are shown.

fore, it is excluded from the interface. Another position for water in 1azp is occupied by the side chain of Arg42 in S7d-D2, which is rotated and interacts with the base of T107.

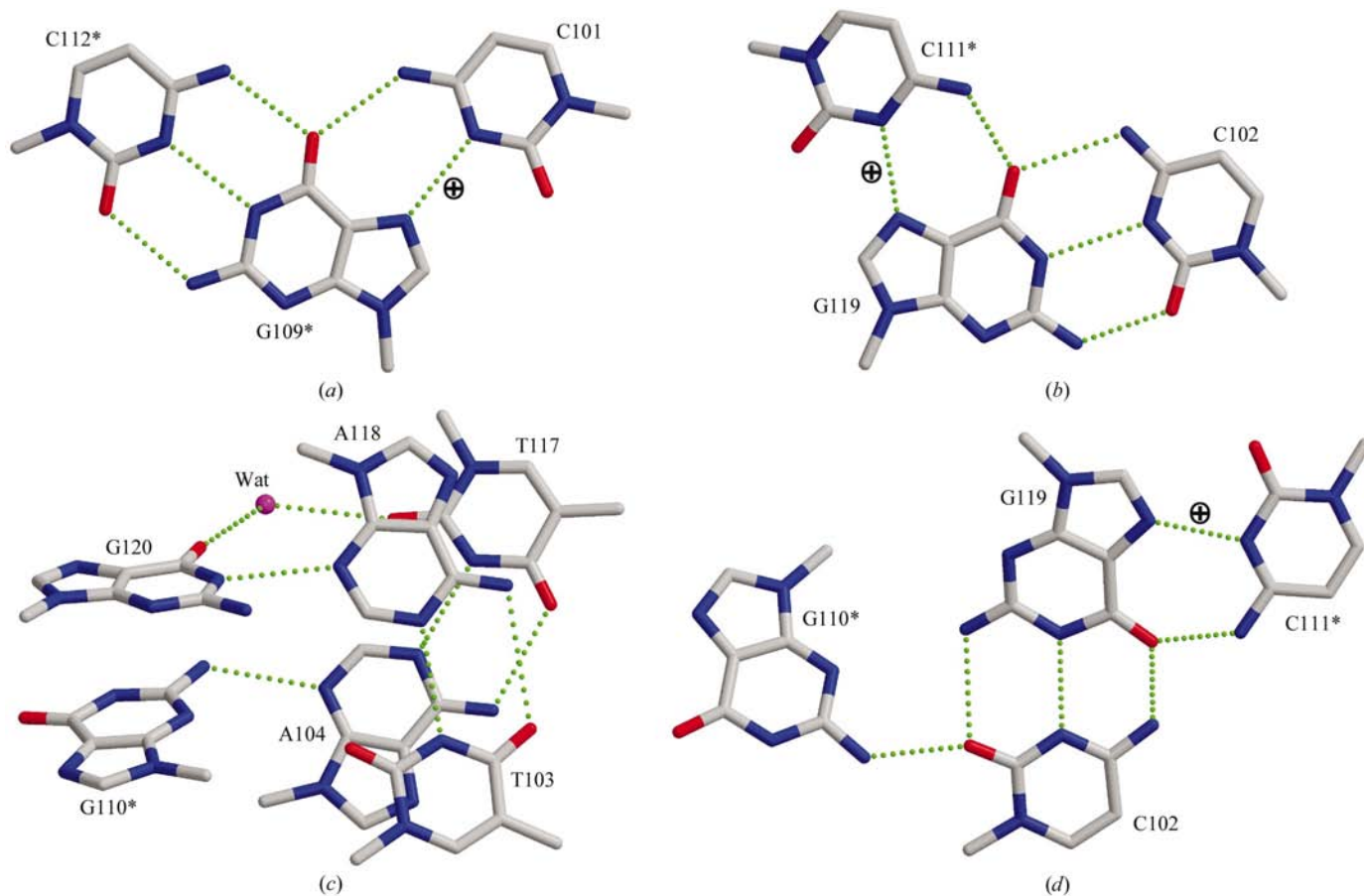
The other three conserved waters are all bound to protein atoms and mediate their interactions with DNA, as shown in Figs. 3(b) and 3(c). Two waters form hydrogen bonds with backbone atoms of Lys7, Ala44 and Ser46, a side-chain atom of Asp49 and a phosphate group of the DNA (G115/A116/G116 in 1azp/S7d-D1/S7d-D2). Additional water-mediated interactions between Sac7d and DNA are also seen, although these are not conserved in all crystal forms. The fourth water is hydrogen bonded to the side chain of Thr33 and surrounded by two lysine side chains (Lys21 and Lys22) and two phosphate groups (T105/A108/A108 and C106/G109/G109). Such conserved interactions, which occur in the nearby region of the intercalation site, may account for the preservation of the overall bending angle of DNA upon binding to Sac7d, regardless of the various single-step kink angles of different DNA sequences.

### 3.3. DNA–DNA interactions and crystal packing

In previous studies of Sac7d–DNA complex structures, it was consistently observed that the DNA octamers form virtual

strands in the crystals. The bases at the 3'-ends are stacked with the bases at the 5'-ends in neighbouring molecules related by the crystallographic  $2_1$  screw axis of symmetry (Robinson *et al.*, 1998; Gao *et al.*, 1998; Su *et al.*, 2000). Similar interactions between the DNA fragments that mimic continuous double helices were also observed in the S7d-D1 and S7d-D2 crystals studied here, as shown in Fig. 4. It appears that the overall packing of the Sac7d–DNA complexes is similar in the crystals of S7d-D1 and S7d-D2, but it is completely different from that of 1azp. The solvent content is about 30% in the former two crystals, but more than 50% in the latter. Interestingly, they are comparable to the crystals of Sso7d–DNA octamer complexes (Gao *et al.*, 1998), which also have smaller unit-cell volumes.

As mentioned above, the nucleotides at both ends of both decamer DNA strands were not involved in regular base pairs. The end-to-end stacking occurs between the base pairs of C102/G119 and G109\*/C112\*, which belong to DNA molecules related by a  $2_1$  screw axis along the crystallographic  $c$  axis (Figs. 4a and 4b). Instead, the terminal nucleotides form triple base pairs, or Hoogsteen structures, with neighbouring DNA molecules *via* non-Watson–Crick hydrogen bonds. Opening of the terminal base pair is common in other DNA



**Figure 5** Interactions between the DNA bases at crystal contacts. (a) and (b) show two CGC\* triple base pairs observed in the S7d-D1 crystal. In (c) the base of G120 is stacked with G110\*. Both are almost perpendicular to and form hydrogen bonds with the two AT base pairs. A mediating water molecule is also shown. (d) shows a similar CGC\* triple base pair as (b) where an additional hydrogen bond is formed between C102 and G110\* in the S7d-D2 crystal. The bonds that require protonated cytosine bases are labelled  $\oplus$ .

crystal structures (Gao *et al.*, 1999). As shown in Figs. 5(a) and 5(b), two such cases are seen in the S7d-D1 crystal. The base of G109\* forms two hydrogen bonds with C101, as does C111\* with G119. In addition, G110\* of the opened 3'-end is base-stacked with G120 of another strand and they form hydrogen bonds with A104 and A118 (Fig. 5c). In the S7d-D2 crystal, the bases of G110\* and C111\* flank the base pair of C102/G119, where three hydrogen bonds are formed (Fig. 5d). The invisible C101 and G120 are probably disordered in this crystal owing to lack of base-pair and lattice-contact interactions. The cytosine bases have to be protonated at N3 ( $pK_a = 4.2$ ) to make the additional hydrogen bonds in Figs. 5(a), 5(b) and 5(d). The 50 mM  $KH_2PO_4$  used in growing S7d-D2 crystals gave a pH value of about 4.0 without adding the conjugate base  $K_2HPO_4$ . The pH in the S7d-D1 crystal might also become less than 6.8 by exposure to air, as it was not within the effective range of Tris buffer.

In addition to the triple base pairs, the DNA molecules also interact with the Sac7d proteins through crystal contacts. There are also significant interactions between symmetry-related protein molecules. In the S7d-D1 crystal, the most extensive occur between molecules related by a  $2_1$  screw axis along the crystallographic *b* axis (Fig. 4c). This type of interface involves at least 12 and 15 residues, including five nucleotides on one side and two on the other. Three positively charged side chains of Lys5, Lys13 and Lys9 interact with the negative charged phosphate groups and the side chain of T105, G119 and Glu47, respectively. The third type of crystal-contact interactions occurs between molecules related by a  $2_1$  screw axis along the crystallographic *c* axis (Fig. 4a), in which seven and five residues are involved. The side-chain NZ of Lys48 forms hydrogen bonds with the backbone O of Ser18 and Ile20, as does Lys48 O with Lys21 NZ. Molecules related by a unit translation along the crystallographic *a* axis are also in contact, involving five nucleotides on one side and six amino acids on another. The terminal amino group of Met1 interacts with the phosphate group of T113. Two other types of crystal contacts occur between molecules related by  $2_1$  screw axes along the crystallographic *a* axis. The crystal-contact interactions bury about 3700 Å<sup>2</sup> of the total 7700 Å<sup>2</sup> surface areas of the S7d-D1 complex.

In the S7d-D2 crystal, a similar repertoire of interfaces is observed. The second type of interface is not as involved as that of S7d-D1 crystal. Two salt bridges are formed between the side chains of Lys5 and Lys7 and the phosphates of C105 and C106, respectively. In addition, the base of C102 is stacked with the side chain of Arg63. The other types of interface are not very different, except for the last one. In the S7d-D1 crystal this represents only a minor contact between the protein C-terminus and the DNA, but in S7d-D2 crystal it involves four amino acids on one side and five nucleotides on another. The side chain of Arg60 forms a salt-bridge to the phosphate of T107 and the side chain of Glu64 is hydrogen bonded to the base of A108. Furthermore, the side-chain N atoms and the backbone O of Arg63 form hydrogen bonds

with the bases of G109 and C111, respectively. Interestingly, G109 is stacked with C102 of a neighbouring DNA, which is involved in the second type of crystal-contact interaction. The total buried area is 3400 Å<sup>2</sup> of the 7100 Å<sup>2</sup> surface of the S7d-D2 complex.

#### 4. Conclusion

By using longer DNA molecules, we intended to make new Sac7d–DNA complex crystals with two or more protein molecules bound to the same DNA and to investigate how the chromosomal protein molecules interact with their neighbours. However, we observe only one Sac7d bound to one DNA decamer in the new crystal forms, which have similar unit-cell sizes but are not isomorphous. Because of kinking and bending, the DNA molecules form a virtual strand with Sac7d bound to alternate sides and the protein–protein interactions are limited. Another crystal form with similar packing is that of the Sso7d–DNA octamer complex (Gao *et al.*, 1998). In all cases the number of base pairs in a unit length of the virtual DNA strand is eight, perhaps as a result of favoured base stacking at the termini. The ‘extra’ bases in the decamer DNA were opened up and formed Hoogsteen structures. To investigate the protein–protein interactions, further experiments with significant variation in the crystallization conditions, including the length of the DNA, should be carried out.

#### References

- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Choli, T., Wittmann-Liebold, B. & Reinhardt, R. (1988). *J. Biol. Chem.* **263**, 7087–7093.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Edmondson, S. P., Qiu, L. & Shriver, J. W. (1995). *Biochemistry*, **34**, 13289–13304.
- Gao, Y. G., Robinson, H., Sanishvili, R., Joachimiak, A. & Wang, A. H.-J. (1999). *Biochemistry*, **38**, 16452–16460.
- Gao, Y. G., Su, S., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W. & Wang, A. H.-J. (1998). *Nature Struct. Biol.* **5**, 782–786.
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
- McAfee, J. G., Edmondson, S. P., Datta, P. K., Shriver, J. W. & Gupta, R. (1995). *Biochemistry*, **34**, 10063–10077.
- Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* **D50**, 869–873.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Reddy, T. R. & Suryanarayana, T. (1988). *Biochim. Biophys. Acta*, **949**, 87–96.
- Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W. & Wang, A. H.-J. (1998). *Nature (London)*, **392**, 202–205.
- Su, S., Gao, Y. G., Robinson, H., Liaw, Y. C., Edmondson, S. P., Shriver, J. W. & Wang, A. H.-J. (2000). *J. Mol. Biol.* **303**, 395–403.
- White, M. F. & Bell, S. D. (2002). *Trends Genet.* **18**, 621–626.