Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Effect of Rap1 binding on DNA distortion and potassium permanganate hypersensitivity

Yann-Vaï Le Bihan,[a] Béatrice Matot,[a]‡ Olivier Pietrement,[b] Marie-Josèphe Giraud-Panis,[c] Sylvaine Gasparini,[a] Eric Le Cam,[b] Eric Gilson,[c] Bianca Sclavi,[d] Simona Miron[a] and Marie-Hélène Le Du[a]*

[a]CEA/DSV/IBiTec-S/SB2SM, Laboratoire de Biologie Structurale et Radiobiologie, CNRS UMR 8221, University Paris-Sud, Bâtiment 144, CE Saclay, 91191 Gif-sur-Yvette, France, [b]Maintenance des Génomes, Microscopies Moléculaires et Bionanosciences, UMR 8126 CNRS and University Paris Sud, 94805 Villejuif, France, [c]Institute for Research on Cancer and Aging, Nice (IRCAN), University of Nice, CNRS UMR 7284, INSERM U1081, 28 Faculté de Médecine, University of Nice, Nice, France, and [d]LBPA, UMR 8113 du CNRS, ENS Cachan, 61 Avenue du Président Wilson, 94235 Cachan, France

‡ Present address: Laboratoire RMN AIM-CEA, Institut de Myologie – G. H. Pitie-Salpetriere, 83 Boulevard de l'Hopital, 75651 Paris CEDEX 13, France.

Correspondence e-mail: marie-helene.ledu@cea.fr

Received 3 August 2012
Accepted 30 November 2012

PDB Reference: Rap1–DNA complex, 4gfb

Repressor activator protein 1 (Rap1) is an essential factor involved in transcription and telomere stability in the budding yeast *Saccharomyces cerevisiae*. Its interaction with DNA causes hypersensitivity to potassium permanganate, suggesting local DNA melting and/or distortion. In this study, various Rap1–DNA crystal forms were obtained using specifically designed crystal screens. Analysis of the DNA conformation showed that its distortion was not sufficient to explain the permanganate reactivity. However, anomalous data collected at the Mn edge using a Rap1–DNA crystal soaked in potassium permanganate solution indicated that the DNA conformation in the crystal was compatible with interaction with permanganate ions. Sequence-conservation analysis revealed that double-Myb-containing Rap1 proteins all carry a fully conserved Arg580 at a position that may favour interaction with permanganate ions, although it is not involved in the hypersensitive cytosine distortion. Permanganate reactivity assays with wild-type Rap1 and the $Rap1_{[R580A]}$ mutant demonstrated that Arg580 is essential for hypersensitivity. AFM experiments showed that wild-type Rap1 and the $Rap1_{[R580A]}$ mutant interact with DNA over 16 successive binding sites, leading to local DNA stiffening but not to accumulation of the observed local distortion. Therefore, Rap1 may cause permanganate hypersensitivity of DNA by forming a pocket between the reactive cytosine and Arg580, driving the permanganate ion towards the C5–C6 bond of the cytosine.

1. Introduction

Telomeres are specialized nucleoprotein complexes at the end of linear eukaryotic chromosomes that control numerous DNA processes. They protect the linear chromosome extremity against DNA-repair machinery, regulate the transcription of nearby genes, control the terminal replication of chromosomal DNA and localize chromosome ends within the nuclear space (Greider & Blackburn, 1989; Gilson & Géli, 2007; de Lange, 2009). Alteration of telomere structure critically affects these processes, leading to genome instability and the development of many diseases, including cancer (Blackburn, 2000). Telomeric DNA consists of repeated short G-rich sequences with a single-stranded tail on the 3′-oriented strand. Telomeric DNA tracts vary from a few hundred base pairs in lower eukaryotes (ciliates and yeast) to thousands of base pairs in higher eukaryotes (plants and vertebrates) (Blackburn, 1994; Rhodes & Giraldo, 1995). In *Saccharomyces cerevisiae*, the telomere nucleoproteic assembly includes Rap1, its functional

partners Rif1, Rif2, Sir3 and Sir4, and the single-stranded DNA-binding protein Cdc13 and its functional partners Stn1 and Ten1 (Giraud-Panis *et al.*, 2010). Rap1 directly binds double-stranded DNA through a central double-Myb domain (DBD), recruits the Rif and Sir proteins through a C-terminal globular domain (RCT) and contains a single BRCT domain located in its N-terminal part (Morse, 2000). The remaining 40% of the peptide chain forms unstructured regions that confer high flexibility to the molecule. Rap1 tightly wraps DNA through its DBD, leading to orientation of the N- and C-terminal regions on each side of the DNA axis. The N-terminal moiety remains flexible, whereas the DBD-RCT partly gains stability upon DNA binding (Matot *et al.*, 2012). Rap1 binds the 13 bp double-stranded DNA consensus sequence ACACCCRYACAYM, which includes the half-sites ACACC or ACAYC at positions 1–6 and ACATY or ACAYM at positions 9–13 (König *et al.*, 1996; Graham & Chambers, 1994; Conrad *et al.*, 1990). Footprinting experiments coupled with bandshift assays have shown that Rap1 binds successive telomeric repeats simultaneously, with an average frequency of one molecule per 18 bp (Gilson *et al.*, 1993). Each Rap1 binding site is associated with a specific $KMnO_4$ hypersensitive site. This ability of Rap1 to induce $KMnO_4$ hypersensitivity at the specific position ACACCCA**C**ACACC suggests that its interaction may affect the DNA conformation (Gilson *et al.*, 1993). Potassium permanganate is commonly used as a probe to characterize DNA–protein interaction *in vitro* and *in vivo* (Spicuglia *et al.*, 2004; Sclavi, 2008). It oxidizes the C5–C6 bond of pyrimidine bases, primarily thymidine (T), with a preference for nucleotides located on single-stranded DNA (Borowiec *et al.*, 1987). Thus, potassium permanganate is routinely used to detect regions of induced DNA melting. However, in some examples, such as that involving Epstein–Barr virus origin binding protein (EBNA1; Bochkarev *et al.*, 1998) and that involving Rap1 protein, the crystal structure raises further questions regarding this reactivity, as the DNA distortion remains close to the standard B-DNA conformation (König *et al.*, 1996; Matot *et al.*, 2012). It is unknown whether or not the crystal packing constrains the DNA conformation and minimizes the observed distortion. Therefore, it remains unclear whether permanganate hypersensitivity is a consequence of DNA melting or of DNA distortion. Also, a direct protein–permanganate interaction cannot be excluded, although no such interaction has ever been documented. For instance, the successive Rap1 binding sites at telomeres may lead to the accumulation of small distortions that could affect the conformation of telomeric DNA.

In order to better understand the effect of Rap1 binding on DNA conformation and potassium permanganate hypersensitivity, we compared the DNA conformation in various Rap1-DBD–DNA crystal forms. The minor DNA distortion observed encouraged us to check the compatibility of our crystal structure with permanganate interaction by analysing the anomalous signal of manganese observed for crystals soaked in potassium permanganate solutions. An analysis of residue conservation in double-Myb-containing Rap1 molecules incited us to study the effect of a point mutation on potassium permanganate reactivity. These assays revealed that the protein itself participates in this phenomenon. Finally, we performed AFM analysis to visualize the effect of successive Rap1 binding on DNA conformation and report that Rap1 binding is associated with straightening of the DNA molecule but not with accumulation of distortion.

2. Materials and methods

2.1. Crystallization

The cloning, production and purification of His_6-tagged $Rap1_{[358–827]}$, $Rap1_{[1–827]}$ and $Rap1_{[1–827]}$ mutant $Rap1_{[R580A]}$ were performed as described previously (Matot *et al.*, 2012). Based on the crystallization conditions of protein–DNA complexes described in 32 different publications (Supplementary Table S1[1]) and on the use of spermine and cobalt hexammine as described in Nowakowski *et al.* (1999), we designed a 48-experiment screen through an incomplete factorial approach as implemented in *SAmBA* (http://www.igs.cnrs-mrs.fr/samba/; Audic *et al.*, 1997; Supplementary Table S2). Crystallization experiments were performed using the $Rap1_{[358–827]}$ construct in complex with various double-stranded DNA oligonucleotides of between 19 and 31 bp in length with blunt or cohesive ends (Eurogentec). Crystallization trials were performed at PF6 Pasteur Institute using a MatrixMaker robot to prepare our specific crystallization screen and a Cartesian robot to complete the crystallization experiments. Initial crystallization hits were optimized using sitting-drop vapour-diffusion experiments in our laboratory. The most suitable crystals were obtained at room temperature with a $Rap1_{[358–827]}$–DNA complex concentration of between 3.3 and 5.5 mg ml^{-1} in a solution consisting of 100 m*M* Tris–HCl pH 8.0, 20% PEG MME 550, 100 m*M* $CaCl_2$, 5% glycerol. Two different crystal forms were obtained using two different oligonucleotides, named SG19 and SG24 after their respective crystal space groups (Table 1; Matot *et al.*, 2012). For the $KMnO_4$ data set, crystals of $KMnO_4$ were directly introduced into a crystallization drop containing SG19 complex crystals before cooling.

2.2. Data collection and structure determination

Native diffraction data were collected on the PX1 beamline at the SOLEIL synchrotron and were reduced with *XDS* (Kabsch, 2010). The structures were solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using PDB entry 1ign as a model probe (König *et al.*, 1996). Electron density and SDS–PAGE analysis of dissolved crystals revealed that the $Rap1_{[353–827]}$ construct was partly degraded and that in addition to the 31 bp oligonucleotide only the double-Myb moiety and part of the linker between the DBD and the C-terminus was present in the crystals. Both structures were refined using *BUSTER* (Bricogne *et al.*, 2009). Diffraction data for SG19 complex crystals soaked in the presence of $KMnO_4$

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: CB5017). Services for accessing this material are described at the back of the journal.

Table 1
Statistics of data collection, refinement and model validation.

Values in parentheses are for the last shell.

Data set	SG19 (Matot *et al.*, 2012)	SG24	SG19 + $KMnO_4$
Oligonucleotide	5′-ACCT**GGTGT**GTG**GGTGT**TGTGT**GGTGT**TCAC -3′ 3′-GA**CCACA**CAC**CCACA**ACACA**CCACA**AGTGTG -5′	5′-ACCT**GGTGT**GTG**GGTGT**TGTGT**GGTGT**TCAC -3′ 3′-TGTGGA**CCACA**CAC**CCACA**ACACA**CCACA**AG -5′	5′-ACCT**GGTGT**GTG**GGTGT**TGTGT**GGTGT**TCAC -3′ 3′-GA**CCACA**CAC**CCACA**ACACA**CCACA**AGTGTG -5′
Wavelength (Å)	0.98	0.98	1.89
Space group	$P2_12_12_1$	$I2_12_12_1$	$P2_12_12_1$
Diffraction limits (Å)	2.95 (3.13–2.95)	2.99 (3.17–2.99)	4.10 (4.32–4.10)
Unit-cell parameters (Å, °)	a = 40.6, b = 102.9, c = 116.8, $\alpha = \beta = \gamma$ = 90	a = 63.8, b = 122.6, c = 149.4, $\alpha = \beta = \gamma$ = 900	a = 40.6, b = 103.2, c = 115.7, $\alpha = \beta = \gamma$ = 90
R_{merge}	0.138 (0.688)	0.121 (0.701)	0.149 (0.376)
No. of unique reflections	9831	11279	4148
$\langle I/\sigma(I)\rangle$	14.32 (3.39)	16.61 (3.09)	7.0 (2.8)
Completeness (%)	90.1 (68.8)	92.4 (78.7)	99.8 (100)
Molecular replacement	*Phaser*, one solution with 75% threshold	*Phaser*, one solution with 75% threshold	
LLG	1305.76	905	
Refinement	*BUSTER*	*BUSTER*	*BUSTER*
Refinement resolution (Å)	2.95	2.99	4.10
R_{work}	0.1923	0.2044	0.1973
R_{free}	0.2650	0.2690	0.3265
Figure of merit	0.869	0.873	0.928
No. of residues	223	218	
No. of bases	62	60	
No. of water molecules	6	12	
R.m.s.d., bonds (Å)	0.0010	0.0010	0.0110
R.m.s.d., angles (°)	1.48	1.40	1.55
PDB entry	3ukg	4gfb	

were collected on the PX1 beamline at the Mn *K* edge wavelength (1.89 Å), integrated with *MOSFLM* and reduced with *SCALA* from the *CCP*4 package (Winn *et al.*, 2011). Two runs of *BUSTER* were performed to improve the phases before calculating the anomalous difference map. Because of the poor diffraction of the $KMnO_4$-soaked crystals [4.1 Å resolution, $I/\sigma(I)$ = 7.0], the integration step was critical to extract the anomalous signal. We performed successive data-integration runs with the *SCALEIT* software (Howell & Smith, 1992) using different levels of outlier rejection cutoff. For each integrated data set, we calculated the anomalous difference electron-density maps using calculated phases from the refined structure of the Rap1–DNA complex. The quality of the anomalous signal was evaluated by inspection of the anomalous difference peaks using the presence of small peaks typical of the phosphorus anomalous signal along the DNA-molecule backbone as a quality reference. Statistics of data collection, refinement and model validation are shown in Table 1.

2.3. Conservation analysis

We first used *PSI-BLAST* (Altschul *et al.*, 1990) to search for homologues of the $Rap1_{[360–601]}$ sequence and then selected sequences with 80% homology coverage using *Blammer*. The 12 sequences identified were aligned with *Kalign* (Lassmann & Sonnhammer, 2005) and the resulting alignment, together with the SG19 coordinate file, were processed with *ConSurf* (Ashkenazy *et al.*, 2010) for conservation analysis. Residue conservation was visualized using *WebLogo* for the sequence (Crooks *et al.*, 2004) and using *PyMOL* for the three-dimensional structure (DeLano, 2002).

2.4. Permanganate-hypersensitivity assay

$KMnO_4$ footprinting experiments were performed on short double-stranded oligonucleotides of telomeric sequence using either wild-type Rap1 (wtRap1) or the $Rap1_{[R580A]}$ mutant. Adjustments of the $KMnO_4$ hypersensitivity assay were made with respect to previously published assays (Gilson *et al.*, 1993; Idrissi *et al.*, 1998), mostly because shorter oligonucleotides were used. Two 21-base single-stranded oligonucleotides, 5′-GCACACCCACACACCAGGACC-3′ and 5′-TCCTGG-TGTGTGGGTGTGCGG-3′, were purchased from Eurogentec. The C-rich strand was radiolabelled at its 5′-terminus using [γ-^{32}P]dATP and T4-PNK, purified using Bio-Spin 6 columns (Bio-Rad) and annealed with a 1.2-fold molar excess of the complementary G-rich strand. Protein–DNA complexes were formed by mixing 2 n*M* labelled double-stranded oligonucleotide with 40 n*M* of either wtRap1 or $Rap1_{[R580A]}$ in a binding buffer consisting of 20 m*M* Tris pH 7.5, 150 m*M* NaCl. Binding reactions were incubated for 10 min on ice followed by 10 min at room temperature (RT). $KMnO_4$ reactions were achieved by the addition of 18 m*M* $KMnO_4$ followed by 1 min incubation at RT. The reactions were stopped by the addition of 470 m*M* β-mercaptoethanol, 0.5 *M* sodium acetate pH 5.2, 0.03 mg ml^{-1} calf thymus DNA. Proteins were removed by extraction with phenol/chloroform/isoamyl alcohol (50:49:1) and the DNA was precipitated with ethanol and dried. Oxidized oligonucleotides were cleaved by dissolution in 1 *M*

piperidine and incubation at 363 K for 25 min. Samples were dried and subjected to two cycles of dissolution in water and lyophilization. The resulting DNA fragments were dissolved in formamide denaturing buffer, heated for 5 min at 363 K and subjected to denaturing electrophoretic migration in a gel consisting of 20% acrylamide/bisacrylamide (19:1), 7 *M* urea and 1× TBE, with 0.06 *M* sodium acetate pH 5.2 added to the lower reservoir of the electrophoretic apparatus to improve the resolution of the very short fragments (Hsu *et al.*, 2006). After electrophoresis, the gels were dried, exposed for autoradiography and viewed using a Typhoon PhosphorImager (GE Healthcare). A G+A-specific DNA-sequencing reaction was performed on the radiolabelled C-rich strand and loaded onto each gel as a reference to localize the $KMnO_4$-sensitive sites (Maxam & Gilbert, 1977). The intensities of the bands in the gels were quantified using the *ImageQuant* 5.0 software. The following equation, adapted from calculation of fractional saturation of binding sites used in DNase I footprinting assays (Brenowitz *et al.*, 2001), was used to normalize band intensities and gives the normalized $KMnO_4$ sensitivity (NKS),

$$\mathrm{NKS} = \left[\frac{(D_{n,\mathrm{site}}/D_{n,\mathrm{std}})}{(D_{r,\mathrm{site}}/D_{r,\mathrm{std}})}\right] - 1, \qquad (1)$$

where D refers to the integrated optical density of a band, n refers to any lane corresponding to a protein–DNA sample, r refers to the reference lane (control sample with no protein added to the reaction mixture), site refers to the band for which the $KMnO_4$ sensitivity is to be determined and std refers to a band for which the integrated optical density does not change for any samples (with or without $KMnO_4$ or with or without protein).

2.5. ITC experiments

ITC titrations were performed at 283 K in a VP-ITC calorimeter (GE Healthcare). Before measurement, all solutions were degassed under vacuum. The two partners were prepared in the same buffer consisting of 50 m*M* Tris pH 7.5, 150 m*M* NaCl, 10 m*M* β-mercaptoethanol. Wild-type Rap1 and $\mathrm{Rap1}_{[\mathrm{R580A}]}$ (4 μM) were titrated against a 21 bp telomeric DNA fragment (40 μM) at 283 K using automatic 10 µl injections. The thermodynamic parameters ΔH (enthalpy change), n (stoichiometry) and K_a (association constant) were obtained by nonlinear least-squares fitting of the experimental data using the *Origin* software provided with the instrument. Standard equations were used to calculate the free energy of binding (ΔG) and the entropy (ΔS). Because degradation was observed in the crystal obtained using $\mathrm{Rap1}_{[358–827]}$, we performed degradation experiments using the same proteins as in the permanganate assay and ITC experiments. After 4 h at 277 K, no degradation of the proteins was detected by SDS–PAGE, consistent with an effect on the $KMnO_4$ reactivity or on DNA affinity that was exclusively related to the R580A mutation.

2.6. AFM sample preparation

16 Rap1 sites (5′-ACACCCRYACACC-3′) were inserted by iterative cloning into the pUC19 plasmid between the *Bam*HI and *Bgl*II restriction sites. A 735 bp fragment with the 16 Rap1 sites at its centre was obtained by PCR amplification with the primers 5′-TGAGAGTGCACCATATGCGGTGTG-3′ (forward) and 5′-TCTCCCCGCGCGTTGGCCGATTCA-3′ (reverse). The PCR product was purified on an anion-exchange MiniQ column (GE Healthcare) with a SMART system (Amersham Biosciences), precipitated with ethanol and suspended in TE buffer (10 m*M* Tris pH 7.5, 1 m*M* EDTA). The homogeneity of the length of the DNA molecules was verified by gel electrophoresis and by TEM (Dupaigne *et al.*, 2008) and the concentration was determined by measurement of the absorption at 260 nm using a GeneQuant 1300 spectrophotometer (GE Healthcare). The Rap1–DNA complexes were prepared with a 735 bp DNA fragment at 2 n*M* concentration (*i.e.* 32 n*M* of Rap1 binding sites) and Rap1 protein at 64 n*M* concentration (*i.e.* two proteins per DNA-binding site) in 10 m*M* Tris pH 7.6, 100 m*M* NaCl buffer. The solutions were incubated for 15 min at 303 K. 5 µl of DNA or DNA–Rap1 solution was deposited on a freshly cleaved mica surface (muscovite mica V-1 quality, EMS) pretreated with 20 μM spermidine for 1 min (Hamon *et al.*, 2007), incubated for 2 min and rinsed with 25 µl 0.2%(*w*/*v*) uranyl acetate solution (Révet & Fourcade, 1998). The surface was blotted and dried.

2.7. AFM imaging

Imaging was carried out in intermittent contact mode with a MultiMode system operating with a NanoScope V controller (Bruker). We used silicon AC160TS cantilevers (Olympus) with resonance frequencies of about 300 kHz. All images were collected at a scan frequency of 1 Hz and a resolution of 2048 × 2048 pixels. The images were analysed with the *NanoScope* software and a third-order polynomial function was used to remove the background.

2.8. Analysis of DNA curvature

The local DNA curvature of DNA molecules in our model system is characterized by the ratio, S/D, of the curvilinear distance (S) to the linear distance (D) (Muzard *et al.*, 1990) between points A and B, which are located 79 nm (*i.e.* 232 bp) from each DNA end and delimit the Rap1 binding-site region (see Figs. 7*a* and 7*b*). This ratio is close to 1 for a straight region and greater than 1 for a curved region. For each DNA molecule, we determined the position of points A and B using the *ImageJ* software (NIH) and measured the distances S and D. As a control, the total length of all DNA molecules analysed was measured and any molecule with an incorrect length was excluded from the statistical analysis.

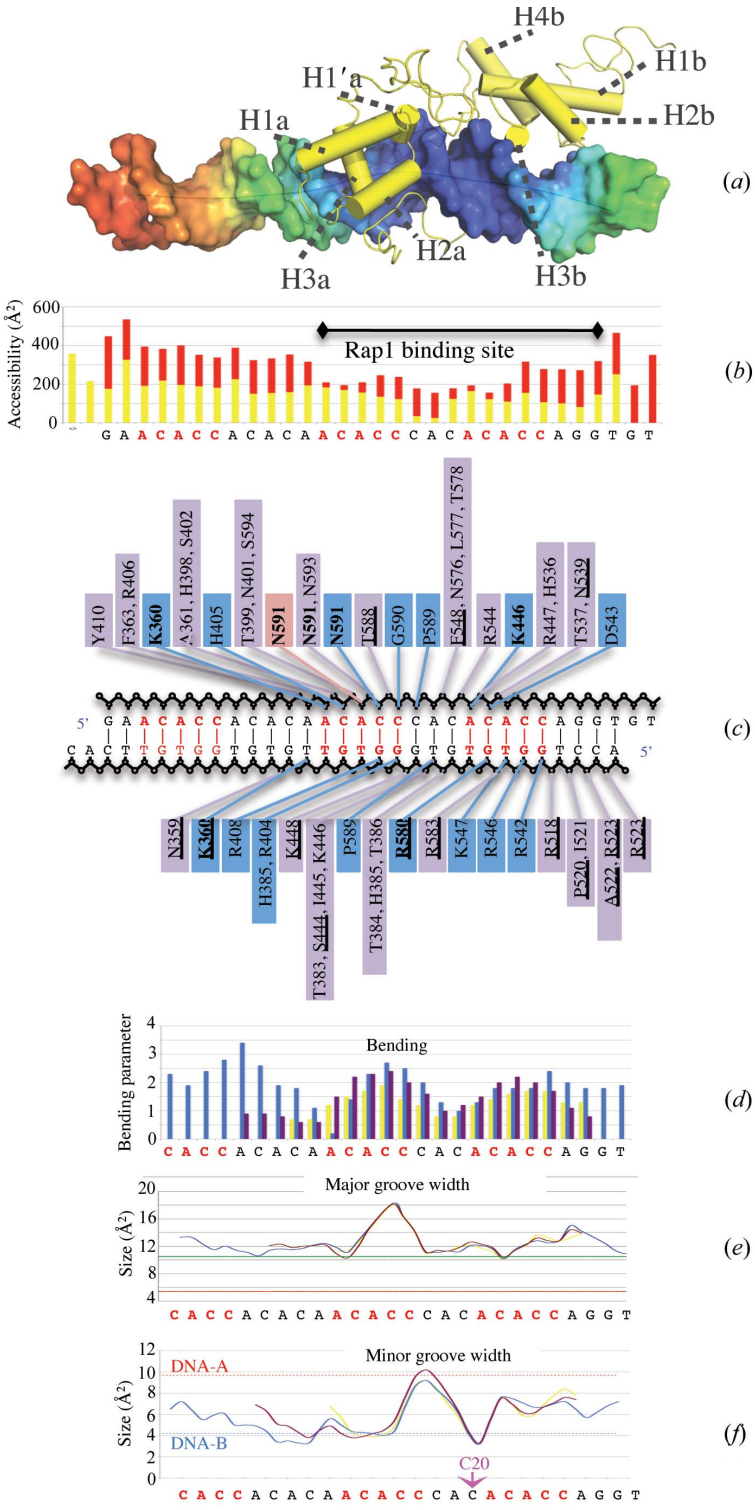


Figure 1
(*a*) Overall structure of Rap1–SG24 with the DNA envelope coloured according to *B* factor; the Rap1$_{[360–602]}$ cartoon is shown in yellow and the bending axis is shown as a dark line. (*b*) Accessibility of C-rich (red) and G-rich (yellow) strands. (*c*) Diagram of the interaction of RAP1 with phosphate (mauve), sugar (pink) or base (blue). Residues in bold are engaged in multiple interactions and underlined residues do not interact with DNA in the 1ign structure. (*d*–*f*) DNA-conformation parameters of SG19 (blue), SG24 (red) and 1IGN (yellow), as measured with the *Curves*+ software (Lavery *et al.*, 2009). (*d*) Histogram of DNA bending. (*e*) Curve of major-groove width. (*f*) Curve of minor-groove width. The pink arrow indicates the hypersensitive Cyt20. The C-rich strand sequence is provided as a reference in each panel.

3. Results and discussion

3.1. Structure analysis

We previously solved the X-ray structure of Rap1$_{[353–602]}$ in complex with a double-stranded DNA oligonucleotide that contains three telomeric hemi-sites (Matot *et al.*, 2012). Using our specifically designed crystallization screening, we obtained a second crystal form that diffracted to 3.0 Å resolution using different overhang extremities of the oligonucleotide (Table 1). The two crystal forms are called SG19 and SG24 after their respective crystal space groups. Superimposition of our structures with the available Rap1–DNA complex structure (PDB entry 1ign; König *et al.*, 1996) shows r.m.s. deviations of 1.77 Å for SG19 and 1.51 Å for SG24. In both structures Rap1$_{[353–602]}$ binds exclusively to two hemi-sites separated by three nucleotides; the additional DNA region, containing one hemi-site and five nucleotides, remains free (Figs. 1*a* and 1*b*).

Our structures provide useful new information regarding the protein, the DNA and the protein–DNA and protein interdomain interactions. Three protein regions were missing in the previous 1ign structure: loops 482–512, 565–571 and 579–586. In SG19 and SG24 the loop corresponding to 482–512 is better defined, although residues 480–498 and 483–501, respectively, were not visible in the electron-density map. We observed no secondary-structure elements in this loop, which adopts slightly different conformations in the two crystal forms, with no particular interaction network. The second loop (residues 565–571) is well defined in SG19 and lacks one residue in SG24 (Supplementary Fig. S1*a*). This region lengthens helix 4 of Myb2 by one helical turn and shows strong backbone-to-backbone interactions between Glu564 O and Arg569 N and between Lys565 O and Asn568 N (Supplementary Fig. S1*b*). These interactions orient the C-terminal part of Myb2 towards the DNA molecule. Beyond this turn, the loop follows helix 4 in an antiparallel manner and is stabilized by hydrogen bonds between the Glu572 backbone and Tyr560 OH and Tyr557 OH. Hydrophobic interactions between Met574 and Leu577 on one side and Tyr557, Tyr552 and Phe548 on the other further stabilize the loop (Supplementary Fig. S1*c*). This hydrophobic network is associated with tight interaction between the backbone N atoms of Asn576, Leu577 and Thr578 and the DNA phosphate of Cyt20 (chain *C*, SG24 numbering; see below). Arg580 plunges into the DNA major groove; its guanidinium group is involved in a bidentate interaction with guanidine 8, forms a stacking interaction with thymidine 7 on the G-rich strand and is parallel to, and located 6.7 Å from, the $KMnO_4$ hypersensitive cytosine 20 on the C-rich strand (see below). The following Arg583 interacts more loosely (at a distance

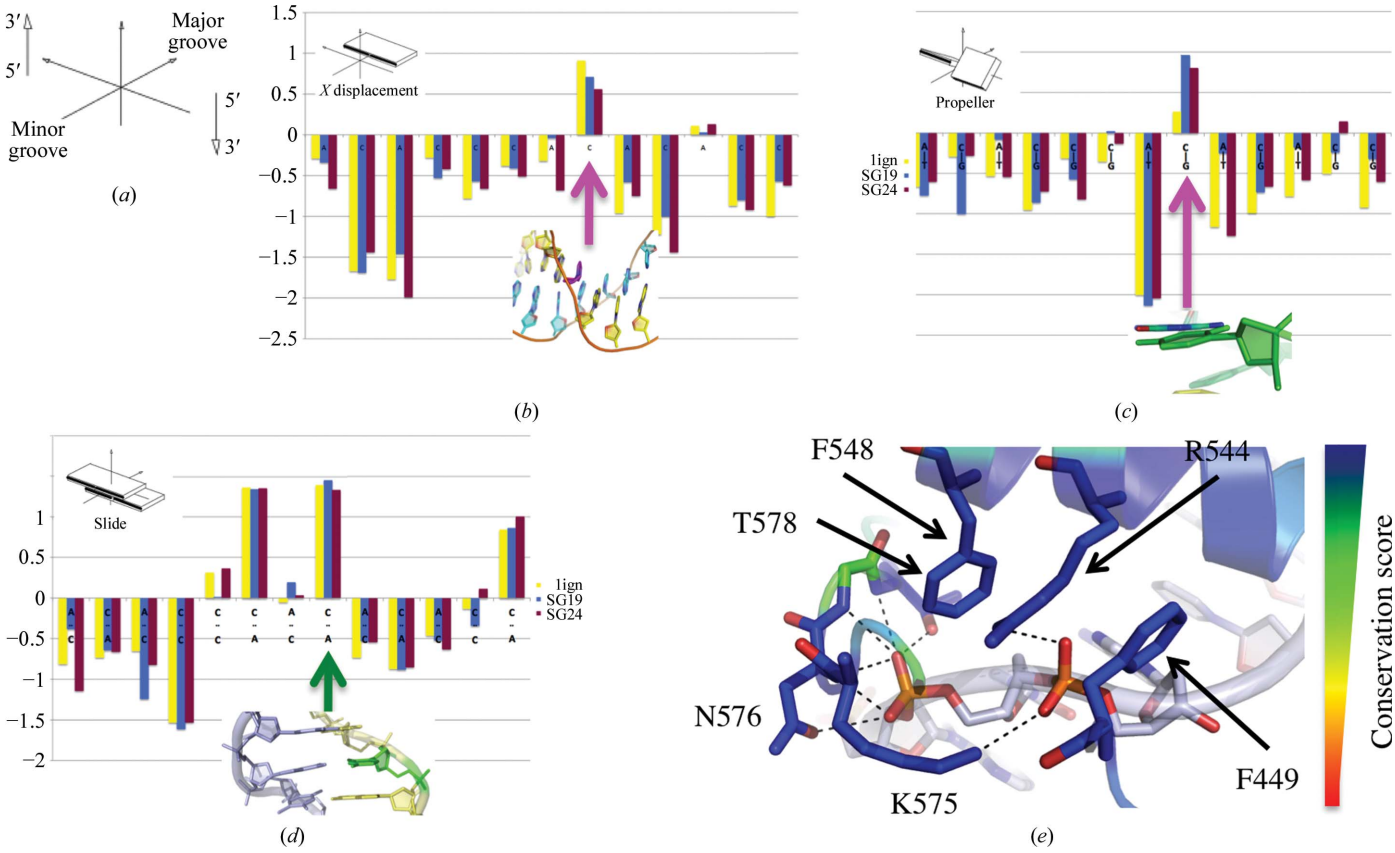


Figure 2
DNA-conformation analysis using the *Curves+* software (Lavery *et al.*, 2009). (*a*) Axis nomenclature; (*b*) displacement along the minor–major groove axis; (*c*) intra-base-pair DNA propeller; (*d*) inter-base-pair DNA slide. The arrows show the distorted base hypersensitive to $KMnO_4$. (*e*) Protein/DNA–backbone interaction network at Cyt20–Ade21. Residues are coloured according to conservation scores from *ConSurf* as indicated on the right (Ashkenazy *et al.*, 2010).

of 3.5 Å) with the phosphates of thymidine 7 and guanidine 8 on the G-rich strand. The next region that deeply penetrates into the DNA major groove includes Pro589 and Gly590, the backbone of which interacts directly with cytosines 16 and 17, and Asn591, the side chain of which interacts with the phosphate of cytosine 16. The following residues, 592–597, form a hook that grips the C-rich strand of the DNA major groove (Asn593 and Ser594) and Myb1 helix 2 (Tyr592 and Lys597). In total, the buried surface of the loop 568–595 is around 670 Å^2 and that of each Myb domain (regions 360–411 and 446–567) corresponds to 920 Å^2. The C-terminal 591–597 extremity of this wrapping loop is important for the functional integrity of Rap1, whereas Arg580 located at the other extremity is not (Matot *et al.*, 2012).

3.2. DNA conformation analysis

The DNA molecule is well defined in both the SG19 and the SG24 complex crystals, although the use of *BUSTER* was critical to build the regions free of protein, which displayed elevated thermal motion and have two missing nucleotides in the C-rich strand of SG24 (Fig. 1*a*). The DNA length is 110 Å, of which 60 Å is covered by Rap1 (Fig. 1*b*). The mapping of protein–DNA interactions enlarges the binding site by two residues in the 3′ direction and three residues in the 5′ direction on the G-rich strand beyond that described for the reference sequence (Fig. 1*c*), which is in agreement with previous observations (Del Vescovo *et al.*, 2004). Bending, distortion and groove-size parameters were calculated with the *Curves+* program (Lavery *et al.*, 2009) using both the full-length DNA of SG19 and SG24 and the Rap1 binding site (ACACCCACACACC) of PDB entry 1ign. The bending of the molecule is more pronounced in SG19 and SG24 than in PDB entry 1ign, in particular in the region AC**ACCC**ACA-ACACACC which corresponds to the path of the C-terminal loop (Fig. 1*d*). In this region the major groove is widened (Fig. 1*e*), whereas the minor groove is narrowed close to cytosine 20 (SG24 numbering; Fig. 1*f*), which corresponds to the potassium permanganate-hypersensitive site (ACACC-CA**C**ACACC; Vignais & Sentenac, 1989; Gilson *et al.*, 1993). Following the nomenclature of Lavery and coworkers (Fig. 2*a*), distortion of the permanganate-hypersensitive Cyt20 occurs through displacement towards the major groove (Fig. 2*b*) and changes in both its intra-base-pair propeller parameter (Fig. 2*c*) and its inter-base-pair slide parameter (Fig. 2*d*). The DNA backbone phosphates of Cyt20 and Ade21 form hydrogen bonds to the side chains of Arg544, Lys575, Asn576 and Thr578, with Arg544 being stabilized through a stacking interaction that involves Phe449 and Phe548 (Fig. 2*e*). The constraints induced by this interaction network are associated

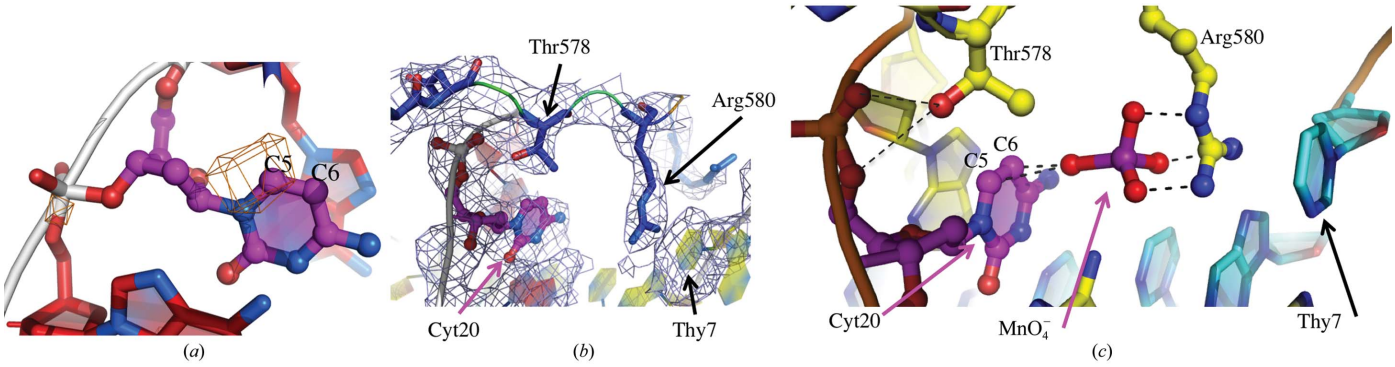


Figure 3
$KMnO_4$-hypersensitive site. (*a*) 2.5σ contour-level anomalous electron-density map at the manganese edge (orange) around Cyt20 (magenta). (*b*) $2F_o - F_c$ electron-density map at the 1σ contour level (blue) of SG19 around Cyt20 (magenta). Residues are coloured according to conservation scores from *ConSurf* as in Fig. 2(*e*). (*c*) Interaction model of the permanganate ion with the guanidinium group of Arg580 and the Cyt20 C5—C6 bond.

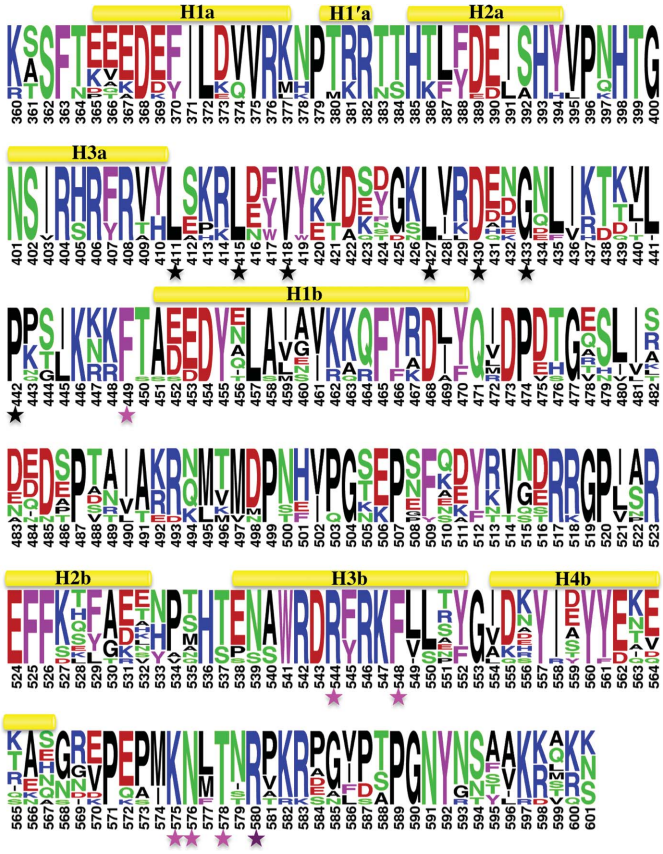


Figure 4
Conservation analysis: residue frequency in 12 double-Myb-containing Rap1 molecules using the *WebLogo* representation (Crooks *et al.*, 2004). The aliphatic residues Gly, Pro, Ala, Val, Leu and Ile and the thiol-containing residues Met and Cys are shown in black, the acidic residues Asp and Glu in red, the basic residues His, Lys and Arg in blue, the aromatic residues Phe, Tyr and Trp in purple and the polar residues Asn, Gln, Thr and Ser in green. The α-helices present in the structure are shown at the top using the same numbering as in Fig. 1(*a*). Residues labelled with stars correspond to conserved residues involved in inter-Myb loop stabilization (black) and in $KMnO_4$-hypersensitive cytosine stabilization (pink) and the conserved Arg580 (purple).

with a distortion of the DNA backbone, in particular at the ε (C4′—C3′—O3′—P) and ζ (C3′—O3′—P—O5′) dihedral angles, for which the values are −97° and 169°, respectively, instead of the 174° and −88° observed in standard B-DNA (Lavery *et al.*, 2009). Two C atoms of the Cyt20 sugar are involved in the distorted dihedral angles, which therefore propagates the distortion up to the base. Potassium permanganate can oxidize pyrimidine bases at the C5—C6 bond, with a preference for thymine, allowing the detection of unpaired or distorted bases (Rubin & Schmid, 1980). It is extensively used to identify protein-binding sites on double-stranded DNA, but very few structural data are available which show minor DNA distortion at the hypersensitive site (Bochkarev *et al.*, 1998). Although we observe a clear distortion at Cyt20 in our crystal structures, its amplitude is small, such that the accessibility of the base in the crystal structures might not be compatible with permanganate interaction.

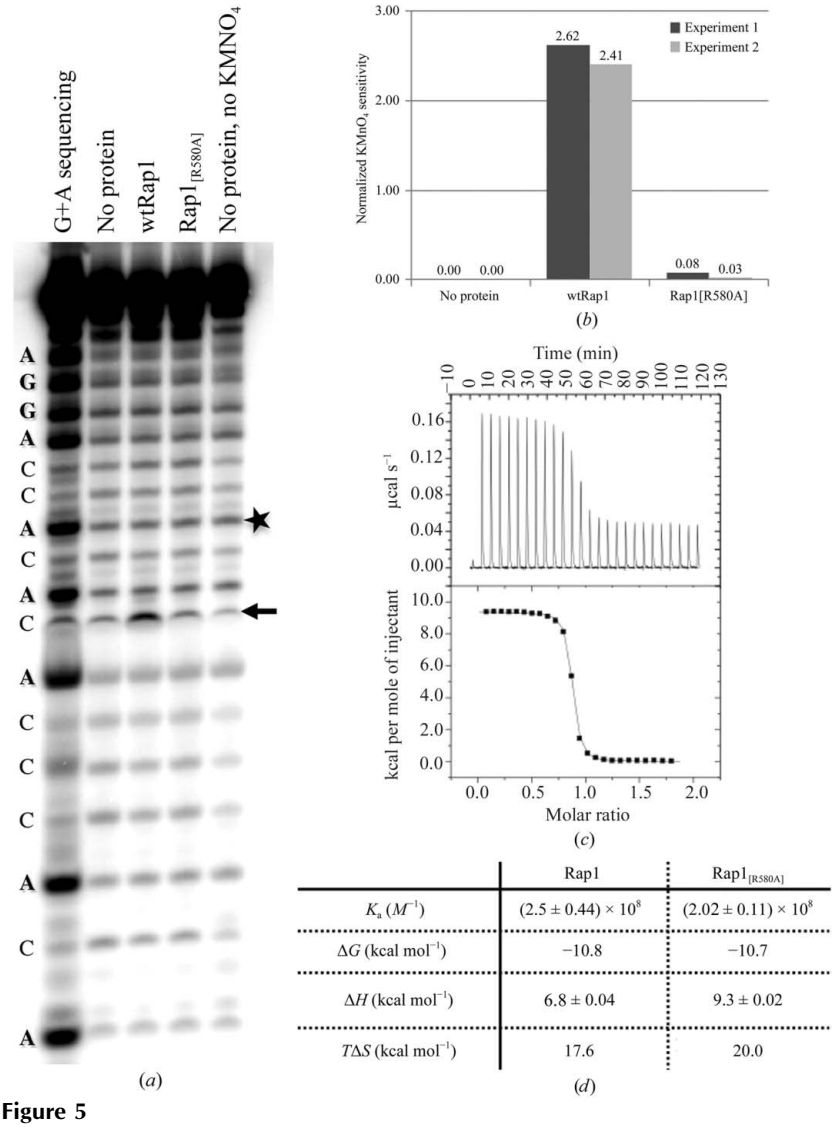
3.3. Permanganate interaction with the Rap1–DNA complex

A truncated form of Rap1 that includes residues 353–598 is sufficient to induce $KMnO_4$ reactivity (Gilson *et al.*, 1993). Therefore, our crystallized complex should allow interaction with $KMnO_4$ if the crystal packing does not constrain the DNA conformation. To test whether this was the case, we collected data at the Mn *K* edge wavelength using an SG19 crystal soaked in the presence of $KMnO_4$. The anomalous difference electron-density map revealed a 3.1σ level peak located close to the C5—C6 bond of the Cyt20 ring that can be associated with a manganese signal (Fig. 3*a*). The low resolution of our data is associated with a ratio of the number of reflections (4148) *versus* the number of parameters to refine (3133) that prevents reliable crystallographic refinement and analysis. However, the presence of an anomalous peak confirms that the DNA conformation in the crystal is compatible with an interaction between one MnO_4^- ion and the C5—C6 bond of Cyt20. Therefore, the small amplitude of the DNA distortion observed in the native crystals suggests that additional determinants may cause the specific hypersensitivity of the cytosine in the sequence ACACCCA**C**ACACC (Cyt20 in the SG24 structure). The position of Arg580 plugged into the DNA major groove and facing the Cyt20

base at a distance of 6.7 Å may implicate this residue as a potential partner (Fig. 3*b*). Indeed, the position and charge of the three N atoms of the guanidinium group of Arg580 are compatible with an interaction with the three double-bonded permanganate O atoms. In agreement with our anomalous data, we manually docked a permanganate ion at this location, which positioned the remaining permanganate O atom with an appropriate orientation and distance (2.5 Å) to react with the C5—C6 bond of Cyt20 (Fig. 3*c*).

3.4. Conservation analysis

Residue conservation during evolution highlights positions that are maintained owing to functional or structural constraints. Conservation analysis of Rap1 was performed using *ConSurf* (Ashkenazy *et al.*, 2010) based on the SG19 coordinate file and 11 sequences of double-Myb containing proteins related to Rap1 (Supplementary Table S3). In addition to the highly conserved Myb domains, other regions display conserved features. Residues involved in the stabilization of the inter-Myb loop are fully conserved among double-Myb Rap1 molecules (Fig. 4), which is in agreement with the conserved positioning of the Myb domains independent of the spacing in the DNA-binding site (König *et al.*, 1996; Taylor *et al.*, 2000; Matot *et al.*, 2012). Similarly, residues Phe449, Arg544, Phe548, Lys575, Asn576 and Thr578, which are involved in Cyt20 interactions and are associated with its distortion, are fully conserved (Figs. 2*e* and 4). This suggests that this distortion is structurally and/or functionally important. Finally, Arg580, which plugs the DNA major groove and faces Cyt20, is also fully conserved (Figs. 3*b* and 4).



	Rap1	Rap1$_{[R580A]}$
K_a (M^{-1})	$(2.5 \pm 0.44) \times 10^8$	$(2.02 \pm 0.11) \times 10^8$
ΔG (kcal mol^{-1})	−10.8	−10.7
ΔH (kcal mol^{-1})	6.8 ± 0.04	9.3 ± 0.02
$T\Delta S$ (kcal mol^{-1})	17.6	20.0

(*d*)

Figure 5
Role of Arg580 of Rap1 in the hypersensitivity of DNA to $KMnO_4$. (*a*) $KMnO_4$ footprinting with no protein (lane 2), with wild-type Rap1 (lane 3; wtRap1), with Rap1$_{[R580A]}$ (lane 4) and with neither protein nor $KMnO_4$ (lane 5). Lane 1, G+A reference sequencing reaction. The arrow shows the position of the known $KMnO_4$-hypersensitive cytosine and the star shows the band used for normalization. (*b*) Normalized quantification results from two independent $KMnO_4$-footprinting experiments. (*c*) ITC experimental curve of Rap1$_{[R580A]}$ with the same double-stranded oligonucleotide as used for $KMnO_4$ footprinting. (*d*) Thermodynamic parameters for the binding of wild-type Rap1 and Rap1$_{[R580A]}$ to DNA as measured by ITC.

3.5. The role of Arg580 in permanganate-induced hypersensitivity

To investigate the molecular events associated with the $KMnO_4$ hypersensitivity of Cyt20, we modified a residue that may favour interaction with $KMnO_4$ but is not involved in nucleotide distortion. Arg580 does not directly interact with Cyt20, with any other nucleotide of the C-rich strand or with residues involved in Cyt20 distortion (Fig. 3*b*); this residue is disordered in the 1ign structure, although the distortion of Cyt20 in this structure is similar to those in our structures (Figs. 2*b*, 2*c* and 2*d*). Arg580 is appropriately oriented to favour interaction with MnO_4^- (Fig. 3*c*) and its location in a loop suggests that its mutation to Ala should not affect the protein structure or its interaction with DNA. Therefore, we constructed an Arg580-to-Ala mutant of Rap1 (Rap1$_{[R580A]}$) and compared the permanganate reactivity of telomeric DNA in the presence of this mutant with that in the presence of wild-type Rap1. Cyt20 in the presence of wild-type Rap1 was clearly reactive, whereas the signal was almost completely absent in the presence of the Rap1$_{[R580A]}$ mutant (Figs. 5*a* and 5*b*). Permanganate reactivity does not necessarily correlate with the affinity of Rap1 for different DNA sequences (Gilson *et al.*, 1993). However, although the assays were performed in the presence of saturating concentrations of Rap1, we have checked that the loss of reactivity was not a consequence of Rap1 affinity since we modified the protein and not the DNA sequence. ITC experiments using Rap1$_{[R580A]}$ led to an association constant of $(2.02 \pm 0.11) \times 10^8\ M^{-1}$ (Fig. 5*c*), which is close to the

$(2.5 \pm 0.44) \times 10^8\ M^{-1}$ reported previously for wild-type Rap1 with the same DNA oligonucleotide (Fig. 5*d*; Matot *et al.*, 2012). Therefore, the observed loss of $KMnO_4$ reactivity in the presence of $Rap1_{[R580A]}$ compared with wild-type Rap1 does not result from weaker binding of $Rap1_{[R580A]}$. Thus, Arg580 is a major determinant of the hypersensitivity of Cyt20 induced by Rap1 binding.

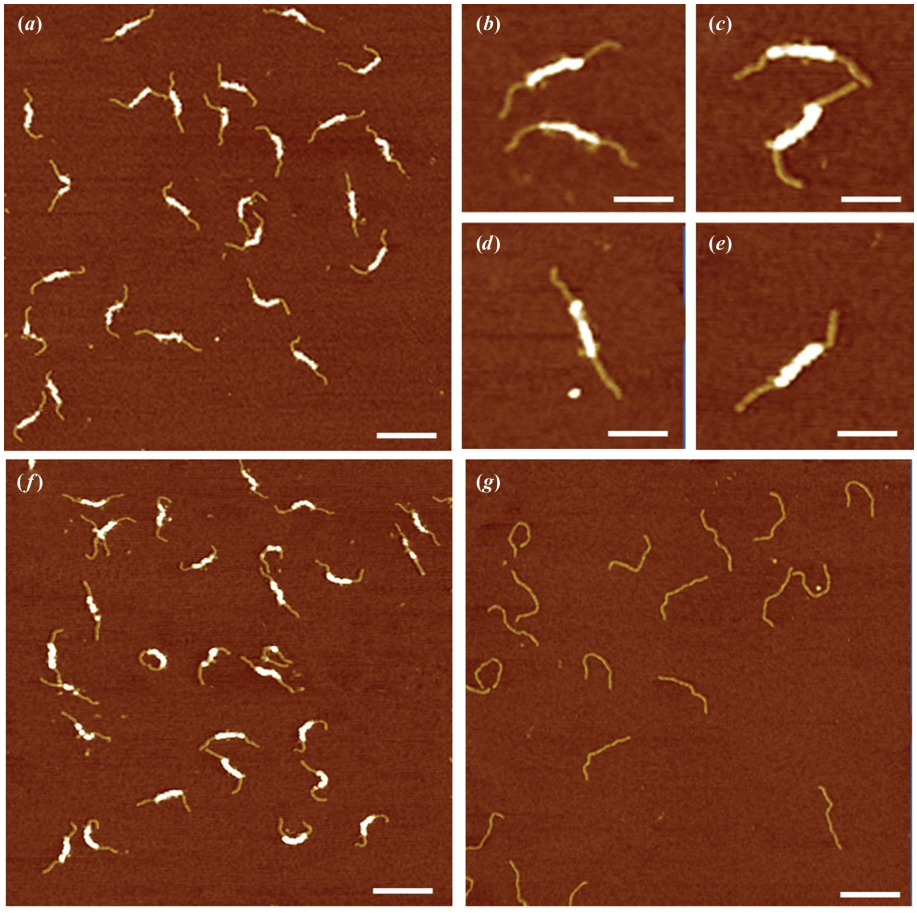


Figure 6
AFM experiments. (*a*) Wild-type Rap1–DNA complexes. (*b*–*e*) Enlarged images of wild-type Rap1–DNA complexes showing local stiffening. (*f*) $Rap1_{[R580A]}$–DNA complexes. (*g*) Bare DNA molecules with 16 Rap1 sites located at the centre of the fragment. The DNA and protein concentrations correspond to a ratio of two proteins per binding site. The scale bars represent 200 nm in (*a*), (*f*) and (*g*) and 100 nm in (*b*–*e*).

3.6. The effect of Rap1 on telomeric DNA stiffness

The number of Rap1 molecules bound to DNA at telomeres varies from one to two molecules at short telomeres to 16 successive molecules at long telomeres. At long telomeres, the weak distortion of DNA associated with Rap1 binding may either be compensated through spacing between Rap1 sites or be amplified along the DNA molecule, leading to a change in the overall DNA conformation. To visualize these possible long-range effects, we performed an AFM experiment using linear double-stranded DNA containing 16 Rap1 sites in complex with Rap1. The length of the DNA molecule was chosen in order to amplify any conformational effect. Wild-type Rap1 (Figs. 6*a*–6*e*) and $Rap1_{[R580A]}$ (Fig. 6*f*) covered the region containing the 16 sites in the DNA, leading to local stiffening; bare DNA molecules did not adopt any particular conformation (Fig. 6*g*). We measured the ratio of the curvilinear (*S*) to the end-to-end (*D*) lengths of the region containing the 16 Rap1 sites (Muzard *et al.*, 1990). A value of 1 for the *S*/*D* ratio reflects a perfectly straight region and higher values reflect greater curvature (Figs. 7*a* and 7*b*). The mean value of the *S*/*D* ratio was

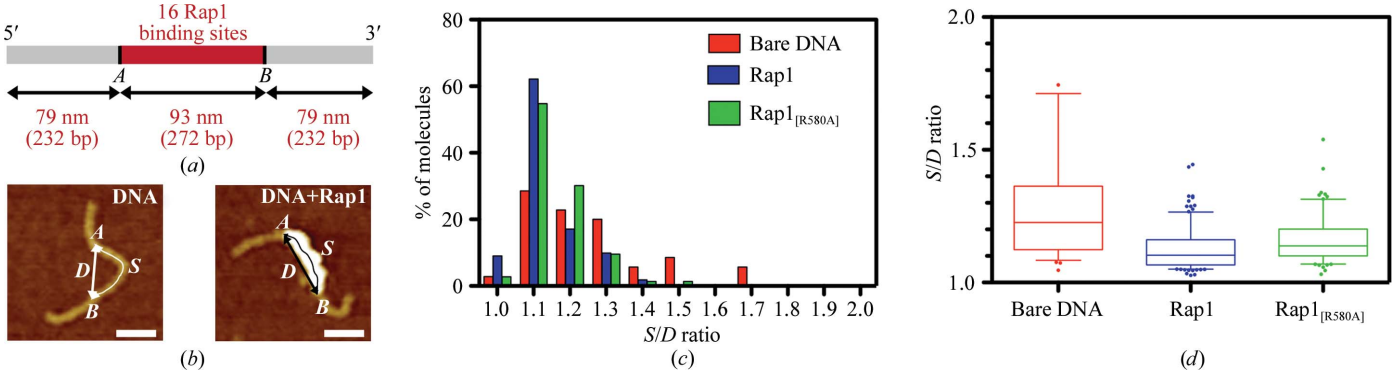


Figure 7
Analysis of DNA curvature. (*a*) Schematic representation of DNA molecules with a central region containing 16 Rap1 binding sites delimited by points *A* and *B* both located 79 nm from the nearest end. (*b*) AFM images of bare DNA molecules and DNA in the presence of Rap1 with illustration of curvilinear (*S*) and direct (*D*) distances. Scale bars represent 50 nm. (*c*) Frequency distribution of the *S*/*D* ratio for bare DNA (n = 137 molecules) and DNA molecules in complex with wild-type Rap1 (n = 128 molecules) or $Rap1_{[R580A]}$ (n = 109 molecules). (*d*) Box–whiskers graph of the frequency distribution of the *S*/*D* ratio for bare DNA and for DNA in complex with wild-type Rap1 or with $Rap1_{[R580A]}$ showing the median value (dashed line) and the 25th and 75th percentiles (rectangle). Data points lower than the 10th percentile and higher than the 90th percentile are represented as circles.

1.32 ± 0.30 for bare DNA and 1.13 ± 0.08 and 1.16 ± 0.10 for the wild-type Rap1 and $Rap1_{[R580A]}$ complexes, respectively (the frequency distribution is presented in Fig. 7*c*). The distributions were significantly different (*P* value lower than 10^{-4}; Mann–Withney *U* test), confirming that binding of either wild-type or mutant Rap1 straightened the DNA molecule (Fig. 7*d*). We therefore conclude that the small distortion of DNA upon Rap1 binding observed in the crystal structures is not related to a particular effect in a telomeric context.

4. Conclusions

We report a structural analysis of the telomeric factor Rap1 in complex with DNA and assess the hypersensitivity to potassium permanganate associated with formation of this complex. Potassium permanganate is routinely used to reveal DNA melting or distortion in DNA–protein complexes (Spicuglia *et al.*, 2004; Sclavi, 2008). We show for the first time that an arginine residue facilitates and probably induces this hypersensitivity. We show that the limited DNA distortion at the hypersensitive site observed in our crystal structures is compatible with an interaction with the manganese ion (Fig. 3*a*), suggesting that the DNA molecule probably adopts a similar conformation in solution. The fact that the sensitivity of Rap1–DNA complexes to $KMnO_4$ is associated with DNA distortion and not DNA melting is in agreement with the similar reactivity observed between linear and negatively supercoiled telomeric DNA (Gilson *et al.*, 1994). The Arg580 residue is fully conserved among all double-Myb proteins related to Rap1, suggesting that it has important functional and/or structural roles. Arg580 plunges into the DNA major groove and faces the hypersensitive cytosine, leaving sufficient space for a permanganate ion as proposed in our manual docking (Fig. 3*c*). We show that an Arg580-to-Ala mutation abolishes the permanganate reactivity of bound DNA, although the affinity of the mutant ($Rap1_{[R580A]}$) for DNA remains similar to that of the wild-type protein (Fig. 5). Also, similar DNA stiffening is caused by the binding of wild-type Rap1 and $Rap1_{[R580A]}$ (Fig. 7). Therefore, we propose that the pocket formed between Cyt20 and Arg580, which can accommodate a negatively charged tetrahedral ion, may be associated with a functional role distinct from DNA distortion. The residues that drive loop 568–580 towards the DNA major groove are also highly conserved (Fig. 4 and Supplementary Fig. S1*c*). This suggests that Arg580 may contribute to the wrapping of Rap1 around the DNA, which is in agreement with general involvement of arginine/major-groove hydrogen bonds in the mechanistic signature of protein–DNA interactions (Garvie & Wolberger, 2001). Finally, our AFM experiments using linear DNA including 16 successive Rap1 sites revealed no accumulation of apparent distortion upon binding of multiple Rap1 molecules (Figs. 6 and 7) that could explain the changes in topology triggered by Rap1 binding (Gilson *et al.*, 1994).

We acknowledge Dr Stéphane Marcand and Dr Rachel Lescasse from the Laboratoire Télomère et Réparation du Chromosome at the Commissariat à l'Energie Atomique, Direction des Sciences du Vivant, Institut de Radiobiologie Cellulaire et Moléculaire for their help with DNA preparation for AFM experiments. We are grateful to Dr Ahmed Haouz and Patrick Weber from PF6 Pasteur Institute for their help and suggestions with crystallization experiments and to the PX1 beamline group at the SOLEIL synchrotron for their assistance during diffraction data acquisition. This work was supported by the Commissariat à l'Energie Atomique, by the Centre National pour la Recherche Scientifique, by the Institut National de la Santé et de la Recherche Médicale, by the Agence Nationale de la Recherche (grant ANR-06-BLAN-0076) and by the Ligue contre le cancer.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Aono, S., Hartsch, T. & Schulze-Gahmen, U. (2003). *Acta Cryst.* D**59**, 576–579.
Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. (2010). *Nucleic Acids Res.* **38**, W529–W533.
Audic, S., Lopez, F., Claverie, J.-M., Poirot, O. & Abergel, C. (1997). *Proteins*, **29**, 252–257.
Blackburn, E. H. (1994). *Cell*, **77**, 621–623.
Blackburn, E. H. (2000). *Nature (London)*, **408**, 53–56.
Bochkarev, A., Bochkareva, E., Frappier, L. & Edwards, A. M. (1998). *J. Mol. Biol.* **284**, 1273–1278.
Borowiec, J. A., Zhang, L., Sasse-Dwight, S. & Gralla, J. D. (1987). *J. Mol. Biol.* **196**, 101–111.
Brenowitz, M., Senear, D. & Kingston, R. (2001). *Curr. Protoc. Mol. Biol.*, Unit 12.14. doi:10.1002/0471142727.mb1204s07.
Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O., Vonrhein, C. & Womack, T. (2009). *BUSTER*. http://www.globalphasing.com.
Conrad, M. N., Wright, J. H., Wolf, A. J. & Zakian, V. A. (1990). *Cell*, **63**, 739–750.
Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004). *Genome Res.* **14**, 1188–1190.
De Lano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Del Vescovo, V., De Sanctis, V., Bianchi, A., Shore, D., Di Mauro, E. & Negri, R. (2004). *J. Mol. Biol.* **338**, 877–893.
Dupaigne, P., Lavelle, C., Justome, A., Lafosse, S., Mirambeau, G., Lipinski, M., Piétrement, O. & Le Cam, E. (2008). *PLoS One*, **3**, e3643.
Escalante, C. R., Shen, L., Escalante, M. C., Brass, A. L., Edwards, T. A., Singh, H. & Aggarwal, A. K. (2002). *J. Struct. Biol.* **139**, 55–59.
Garvie, C. W. & Wolberger, C. (2001). *Mol. Cell*, **8**, 937–946.
Genis, C., Scone, P., Kasahara, H. & Nam, H.-J. (2008). *Acta Cryst.* F**64**, 1079–1082.
Gilson, E. & Géli, V. (2007). *Nature Rev. Mol. Cell Biol.* **8**, 825–838.
Gilson, E., Müller, T., Sogo, J., Laroche, T. & Gasser, S. M. (1994). *Nucleic Acids Res.* **22**, 5310–5320.
Gilson, E., Roberge, M., Giraldo, R., Rhodes, D. & Gasser, S. M. (1993). *J. Mol. Biol.* **231**, 293–310.
Giraud-Panis, M. J., Pisano, S., Poulet, A., Le Du, M.-H. & Gilson, E. (2010). *FEBS Lett.* **584**, 3785–3799.
Graham, I. R. & Chambers, A. (1994). *Nucleic Acids Res.* **22**, 124–130.
Greider, C. W. & Blackburn, E. H. (1989). *Nature (London)*, **337**, 331–337.
Hamon, L., Pastré, D., Dupaigne, P., Le Breton, C., Le Cam, E. & Piétrement, O. (2007). *Nucleic Acids Res.* **35**, e58.
Hirsch, J. A., Wah, D. A., Dorner, L. F., Schildkraut, I. & Aggarwal, A. K. (1997). *FEBS Lett.* **403**, 136–138.
Horton, N. C., Dorner, L. F., Schildkraut, I. & Perona, J. J. (1999). *Acta Cryst.* D**55**, 1943–1945.

Hovde, S., Brooks, A., Strong, K. & Geiger, J. H. (2002). *Acta Cryst.* D**58**, 511–512.
Howell, P. L. & Smith, G. D. (1992). *J. Appl. Cryst.* **25**, 81–86.
Hsu, L. M., Cobb, I. M., Ozmore, J. R., Khoo, M., Nahm, G., Xia, L., Bao, Y. & Ahn, C. (2006). *Biochemistry*, **45**, 8841–8854.
Idrissi, F. Z., Fernández-Larrea, J. B. & Piña, B. (1998). *J. Mol. Biol.* **284**, 925–935.
Imasaki, T., Hashimoto, H., Shimizu, T., Kato, M., Tsuda, J., Kita, K. & Sato, M. (2004). *Acta Cryst.* D**60**, 1165–1166.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kapetaniou, E. G., Kotsifaki, D., Providaki, M., Rina, M., Bouriotis, V. & Kokkinidis, M. (2007). *Acta Cryst.* F**63**, 12–14.
König, P., Giraldo, R., Chapman, L. & Rhodes, D. (1996). *Cell*, **85**, 125–136.
Lange, T. de (2009). *Science*, **326**, 948–952.
Larivière, L., Kurzeck, J., Aschke-Sonnenborn, U., Rüger, W. & Moréra, S. (2002). *Acta Cryst.* D**58**, 1484–1486.
Lassmann, T. & Sonnhammer, E. L. L. (2005). *BMC Bioinformatics*, **6**, 298.
Lavery, R., Moakher, M., Maddocks, J. H., Petkeviciute, D. & Zakrzewska, K. (2009). *Nucleic Acids Res.* **37**, 5917–5929.
Lu, P., Li, Y., Gorman, A. & Chi, Y.-I. (2006). *Acta Cryst.* F**62**, 525–529.
Matot, B., Le Bihan, Y. V., Lescasse, R., Pérez, J., Miron, S., David, G., Castaing, B., Weber, P., Raynal, B., Zinn-Justin, S., Gasparini, S. & Le Du, M.-H. (2012). *Nucleic Acids Res.* **40**, 3197–3207.
Maxam, A. M. & Gilbert, W. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 560–564.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Morse, R. H. (2000). *Trends Genet.* **16**, 51–53.
Muzard, G., Théveny, B. & Révet, B. (1990). *EMBO J.* **9**, 1289–1298.
Ng, C. K. L., Palasingam, P., Venkatachalam, R., Baburajendran, N., Cheng, J., Jauch, R. & Kolatkar, P. R. (2008). *Acta Cryst.* F**64**, 1184–1187.
Nowakowski, J., Shim, P. J., Joyce, G. F. & Stout, C. D. (1999). *Acta Cryst.* D**55**, 1885–1892.
O'Loughlin, T. J., Xu, Q., Kucera, R. B., Dorner, L. F., Sweeney, S., Schildkraut, I. & Guo, H.-C. (2000). *Acta Cryst.* D**56**, 1652–1655.
Pereira, J. H., Ha, S. C. & Kim, S.-H. (2008). *Acta Cryst.* F**64**, 175–178.
Pereira de Jésus, K., Serre, L., Hervouet, N., Bouckson-Castaing, V., Zelwer, C. & Castaing, B. (2002). *Acta Cryst.* D**58**, 679–682.
Pio, F., Ni, C.-Z., Mitchell, R. S., Knight, J., McKercher, S., Klemsz, M., Lombardo, A., Maki, R. A. & Ely, K. R. (1995). *J. Biol. Chem.* **270**, 24258–24263.
Redondo, P., Prieto, J., Ramos, E., Blanco, F. J. & Montoya, G. (2007). *Acta Cryst.* F**63**, 1017–1020.
Révet, B. & Fourcade, A. (1998). *Nucleic Acids Res.* **26**, 2092–2097.
Rhodes, D. & Giraldo, R. (1995). *Curr. Opin. Struct. Biol.* **5**, 311–322.
Richardson, J. M., Finnegan, D. J. & Walkinshaw, M. D. (2007). *Acta Cryst.* F**63**, 434–437.
Rubin, C. M. & Schmid, C. W. (1980). *Nucleic Acids Res.* **8**, 4613–4619.
Sam, M. D., Abbani, M. A., Cascio, D., Johnson, R. C. & Clubb, R. T. (2006). *Acta Cryst.* F**62**, 825–828.
Sam, M. D., Cascio, D., Johnson, R. & Clubb, R. T. (2003). *Acta Cryst.* D**59**, 1238–1240.
Schwartz, T., Shafer, K., Lowenhaupt, K., Hanlon, G., Herbert, A. & Rich, A. (1999). *Acta Cryst.* D**55**, 1362–1364.
Sclavi, B. (2008). *Biochem. Soc. Trans.* **36**, 745–748.
Shanmuganatham, K. K., Ravichandran, M., Howe, M. M. & Park, H.-W. (2007). *Acta Cryst.* F**63**, 620–623.
Spicuglia, S., Kumar, S., Chasson, L., Payet-Bornet, D. & Ferrier, P. (2004). *J. Biochem. Biophys. Methods*, **59**, 189–194.
Suwa, Y., Nakamura, T., Toma, S., Ikemizu, S., Kai, H. & Yamagata, Y. (2008). *Acta Cryst.* F**64**, 171–174.
Tahirov, T. H., Inoue-Bungo, T., Sasaki, M., Fujikawa, A., Kimura, K., Sato, K., Adachi, S., Kamiya, N. & Ogata, K. (2001). *Acta Cryst.* D**57**, 854–856.
Tamulaitiene, G., Grazulis, S., Janulaitis, A., Janowski, R., Bujacz, G. & Jaskolski, M. (2004). *Biochim. Biophys. Acta*, **1698**, 251–254.
Tan, S., Hunziker, Y., Pellegrini, L. & Richmond, T. J. (2000). *J. Mol. Biol.* **297**, 947–959.
Taylor, H., O'Reilly, M., Leslie, A. & Rhodes, D. (2000). *J. Mol. Biol.* **303**, 693–707.
Textor, L. C., Wilmanns, M. & Holton, S. J. (2007). *Acta Cryst.* F**63**, 657–661.
Vassylyeva, M. N., Svetlov, V., Klyuyev, S., Devedjiev, Y. D., Artsimovitch, I. & Vassylyev, D. G. (2006). *Acta Cryst.* F**62**, 1027–1030.
Viadiu, H., Kucera, R., Schildkraut, I. & Aggarwal, A. K. (2000). *J. Struct. Biol.* **130**, 81–85.
Viadiu, H., Vanamee, É. S., Jacobson, E. M., Schildkraut, I. & Aggarwal, A. K. (2003). *Acta Cryst.* D**59**, 1493–1495.
Vignais, M. L. & Sentenac, A. (1989). *J. Biol. Chem.* **264**, 8463–8466.
Vivian, J. P., Porter, C., Wilce, J. A. & Wilce, M. C. J. (2006). *Acta Cryst.* F**62**, 1104–1107.
Vivian, J. P., Wilce, J. A., Hastings, A. F. & Wilce, M. C. J. (2001). *Acta Cryst.* D**57**, 421–424.
Watanabe, S., Kita, A., Kobayashi, K., Takahashi, Y. & Miki, K. (2006). *Acta Cryst.* F**62**, 1275–1277.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.