Published online 13 May 2013 *Nucleic Acids Research, 2013, Vol. 41, No. 13* ***6761–6773***
doi:10.1093/nar/gkt395

Crystal structure and functional insights into uracil-DNA glycosylase inhibition by phage ϕ29 DNA mimic protein p56

José Ignacio Baños-Sanz[1], Laura Mojardín[2], Julia Sanz-Aparicio[1], José M. Lázaro[2], Laurentino Villar[2], Gemma Serrano-Heras[3], Beatriz González[1,*] and Margarita Salas[2,*]

[1]Departamento de Cristalografía y Biología Estructural, Instituto de Química-Física 'Rocasolano' (CSIC), Serrano 119, 28006 Madrid, Spain, [2]Instituto de Biología Molecular 'Eladio Viñuela' (CSIC), Centro de Biología Molecular 'Severo Ochoa' (CSIC-UAM), C/ Nicolás Cabrera 1, Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain and [3]Experimental Research Unit, General University Hospital of Albacete, 02006 Albacete, Spain

Received March 20, 2013; Revised April 17, 2013; Accepted April 18, 2013

ABSTRACT

Uracil-DNA glycosylase (UDG) is a key repair enzyme responsible for removing uracil residues from DNA. Interestingly, UDG is the only enzyme known to be inhibited by two different DNA mimic proteins: p56 encoded by the *Bacillus subtilis* phage ϕ29 and the well-characterized protein Ugi encoded by the *B. subtilis* phage PBS1/PBS2. Atomic-resolution crystal structures of the *B. subtilis* UDG both free and in complex with p56, combined with site-directed mutagenesis analysis, allowed us to identify the key amino acid residues required for enzyme activity, DNA binding and complex formation. An important requirement for complex formation is the recognition carried out by p56 of the protruding Phe191 residue from *B. subtilis* UDG, whose side-chain is inserted into the DNA minor groove to replace the flipped-out uracil. A comparative analysis of both p56 and Ugi inhibitors enabled us to identify their common and distinctive features. Thereby, our results provide an insight into how two DNA mimic proteins with different structural and biochemical properties are able to specifically block the DNA-binding domain of the same enzyme.

INTRODUCTION

Genomic DNA is continuously exposed to damage by internal or external agents, which can generate a variety of DNA lesions threatening genome integrity and cell viability. To prevent the deleterious effects caused by DNA damage, organisms have developed a number of DNA repair mechanisms (1–3). Uracil, a base normally found in RNA, is one of the most frequent lesions in genomic DNA. Uracil may arise in DNA either by misincorporation of deoxyuridine monophosphate (dUMP) instead of deoxythymidine monophosphate (dTMP) during DNA synthesis or by spontaneous deamination of cytosine in DNA. Most DNA polymerases are able to incorporate dUMP and dTMP with similar efficiency (4,5) producing U:A pairs that are not directly mutagenic, but may become genotoxic by impeding sequence recognition carried out by regulatory proteins (6). On the other hand, cytosine deamination can lead to GC→AT transition mutations after the next round of replication posing a serious threat to genome integrity (7,8).

Uracil-DNA glycosylases (UDGs) are the enzymes responsible for removing uracil residues from DNA. UDGs initiate the base excision repair pathway by hydrolysing the N-glycosidic bond between the uracil residue and the deoxyribose sugar of the DNA backbone generating an apurinic-apyrimidinic site (9). UDGs have been classified into four distinct families (10). Members of Family-1 are ubiquitous UDG proteins that are able to excise uracil in both single- and double-stranded DNA (10).

Several structural studies of UDG in complex with DNA have led to propose a complex mechanism of action for this enzyme (11–16). These analyses revealed that UDGs bind, kink and compress the DNA backbone via the action of highly conserved Ser-Pro loops (LII, LIV and LV) while scanning the minor groove for a uracil lesion. The enzyme is hypothesized

*To whom correspondence should be addressed. Tel: +34 91 1964675; Fax: +34 91 1964420; Email: msalas@cbm.uam.es
Correspondence may also be addressed to Beatriz González. Tel: +34 91 5619400; Fax: +34 91 5642431; Email: xbeatriz@iqfr.csic.es

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

to induce a further compression of the DNA backbone flanking the uracil residue, resulting in the flipping of the uracil out of the DNA helix and into the active site pocket of UDG. The hydrophobic side chain of a conserved leucine, Leu272 in human UDG (*Hs*UDG) and Leu191 in *Escherichia coli* UDG (*Ec*UDG), is inserted into the site vacated by the flipped-out uracil. This protruding leucine partially restores base-stacking interactions and probably functions as a 'door stop' for preventing the return of the flipped-out uracil into the DNA helix, as suggested (17).

The binding of UDG to uracil is proposed to induce electron flow from O4′ of the deoxyribose to O2 of uracil. This negative charge is enzymatically stabilized by a neutral conserved histidine (H187 in *Ec*UDG, H268 in *Hs*UDG). A weakly nucleophilic water molecule completes hydrolysis. Protonation on N1 following bond cleavage gives the more stable amide and completes the reaction. As a result, a free uracil molecule and an abasic site in the DNA are obtained (6,14).

Despite the importance of removing uracil from DNA in the maintenance of genome stability, some proteins are known to naturally inhibit the repair activity of UDGs. Particularly, uracil-containing DNA bacteriophages have developed specialized strategies to counteract the cellular uracil excision repair pathway to survive and replicate in their hosts. For instance, phages PBS1 and PBS2 encode an inhibitor of *Bacillus subtilis* UDG (*Bs*UDG) called Ugi, which is critical to preserve the uracil residues incorporated into the uracil-containing phage DNA (18). Ugi is a small (84 amino acids) highly acidic protein that is able to form a physiologically irreversible complex with a variety of UDG proteins in 1:1 molar stoichiometry by mimicking enzyme–DNA interactions (19–22).

Recently, we have reported the identification of a novel inhibitor encoded by the *B. subtilis* phage ϕ29, called protein p56, which is able to bind and block the host UDG activity (23). Although the ϕ29 genome does not contain uracil residues, p56, a small (56 amino acids) highly acidic protein has been proposed to prevent the deleterious effects caused by the host UDG activity in the ϕ29 genome integrity, if uracil is removed from the replicative intermediates. The capacity of p56 for blocking the DNA-binding domain of UDG and the ability of Ugi to replace p56 previously bound to UDG suggests that p56 is able to inhibit UDG by mimicking DNA properties (24).

Bacteriophage proteins p56 and Ugi have likely evolved to successfully inhibit the same enzyme for preserving virus viability. Currently, only a few proteins with DNA mimic features have been characterized (21). Furthermore, the capacity of these proteins to show an effective mimicry of the interactions displayed by DNA with their targets without resulting in cross-reactivity with other DNA-binding enzymes remains elusive.

In this study, we sought to explore the mechanisms underlying *Bs*UDG inhibition by p56. The data obtained from the atomic-resolution crystal structure of *Bs*UDG in complex with p56, together with the results of site-directed mutagenesis analysis, allowed us to determine the most relevant residues involved in this interaction. Thus, our results reveal new structural and functional insights into the mechanism of *Bs*UDG inhibition by p56 and provide additional evidence for understanding the interaction of DNA mimic proteins with their targets.

MATERIALS AND METHODS

DNA substrates

Oligonucleotide ssDNA-U16 (5′-CTGCAGCTGATGCG CUGTACGGATCCCC-GGGTAC-3′) containing a single uracil residue at position 16 was purified electrophoretically on 8 M urea/20% polyacrylamide gels and then 5′-end labelled using [γ-^{32}P]-adenosine triphosphate (3000 Ci mmol-1) (Perkin-Elmer Life Science) and T4 polynucleotide kinase (New England Biolabs). To generate dsDNA substrates, the 5′-^{32}P-labelled ssDNA-U16 oligonucleotides were annealed to complementary non-labelled oligonucleotides (34 mer), containing a guanine residue opposite to uracil. Hybridizations were performed in the presence of 60 mM Tris–HCl (pH 7.5) and 0.2 M NaCl for 10 min at 70°C and then slowly cooled to room temperature. Oligonucleotides were obtained from Isogen Bioscience BV.

Site-directed mutagenesis of *Bs*UDG, GST-*Bs*UDG and p56

Protein mutants were obtained by using the QuickChange site-directed mutagenesis kit provided by Stratagene. To obtain *Bs*UDG mutants, plasmid pT7-4-UDG-wt was used as a template for the mutagenesis reaction with specific oligonucleotides. The PCR reaction was performed using PfuTurboTM DNA polymerase, and then the product was treated with DpnI endonuclease. The amplified DNA was transformed into *E. coli* BL21 (DE3) competent cells. The presence of the desired mutation was confirmed by sequencing the entire gene. The same protocol was performed to generate mutants of Glutathione S-Transferase (GST)-*Bs*UDG (GST was attached to the N-terminal region of *Bs*UDG) and of p56 except for the template vector: pGEX-2T-UDG-wt for GST-*Bs*UDG mutants and pT7-3-p56-wt for p56 mutants. Both plasmids were previously described (25). The specific oligonucleotides used in the mutagenesis reactions are listed in Supplementary Table S1.

Determination of UDG activity

UDG activity was determined as described (25). To test the *Bs*UDG inhibition by p56, the minimal *Bs*UDG amount needed to obtain 50% cleavage of substrate was incubated with the indicated amount of p56 proteins at 4°C for 15 min, and then added to a reaction buffer containing ^{32}P-labelled substrate.

DNA gel retardation assay

The DNA-binding ability of the GST-*Bs*UDG wild-type and the mutants were examined by gel retardation assays. Proteins were incubated with 34 bp 5′-^{32}P-labelled dsDNA U:G (0.7 nM) in a final volume of 20 µl of binding buffer [50 mM Tris–HCl (pH 7.5), 4% (vol/vol) glycerol, 1 mM dithiothreitol (DTT), 0.1 mg/ml of bovine serum albumin]

for 5 min at 4°C. Then the samples were analysed by non-denaturing polyacrylamide gel electrophoresis (6% polyacrylamide), which was pre-run at 30 mA for 20 min. Electrophoresis was carried out at 4°C in TAE buffer (40 mM Tris, 1 mM EDTA, pH set to 7.5 with acetic acid) at 30 mA for 30 min. Gels were vacuum dried, and the radioactive bands were detected by autoradiography and quantified by densitometry.

GST pull-down assay

The pull-down assay was used to determine the interaction between GST-*Bs*UDG and p56 proteins. A mixture containing 30 μg of GST-*Bs*UDG (or GST as a control) and 3 μg of p56 in a final volume of 100 μl of buffer PBS [137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4 (pH 7.4)] was incubated for 15 min on ice (all the following steps were carried out at 4°C). After that, the samples were mixed gently with 60 μl of a 50% slurry of glutathione-sepharose beads (Glutathione Sepharose 4 Fast Flow from GE Healthcare), and then 700 μl of PBS was added to ensure an efficient mixing of reagents. The samples were incubated overnight with an end-over-end mixing and then centrifuged at 2000 rpm for 5 min. The beads were washed once with 10 volumes of PBS and mixed end-over-end during 20 min and then once more with TEN600 buffer [20 mM Tris–HCl (pH 7.5), 0.1 mM EDTA and 600 mM NaCl] to remove unbound proteins. The elution of the GST-*Bs*UDG and the bound p56 proteins was performed by boiling the samples in loading buffer [37 mM Tris–HCl (pH 6.8), 2% (w/vol) sodium dodecyl sulfate, 4% (vol/vol) β-mercaptoethanol and 13% (vol/vol) glycerol] during 5 min. Samples were analysed by 16% (w/vol) polyacrylamide Tris–Tricine gels and visualized using Coomassie blue staining.

Crystallization of *Bs*UDG-p56 complex and *Bs*UDG

For *Bs*UDG-p56 complex crystallization, ammonium sulphate pellets of UDG from *B. subtilis* and p56 from ϕ29 were resuspended in buffer A [20 mM Tris–HCl (pH 7.5) and 1 mM EDTA], mixed in a 1:1.5 molar ratio to a final concentration of 11.5 mg/ml and dialyzed against buffer A. Inital crystallization trials were done using a NanoDrop robot (Innovadyne Technologies Inc.). Two commercial screens, Crystal Screen I (Hampton Research) and JCSG+ screen (Qiagen), yielded preliminary crystals in different conditions that were further optimized. The best crystals obtained for *Bs*UDG-p56 complex appeared at 4°C after 7–8 days in sitting drops prepared by mixing 1 μl of the protein complex plus 1 μl of reservoir solution and using 19% PEG 8000, 0.1 M Tris–HCl (pH 8.5), 0.25 M $MgCl_2$ as precipitant. The crystals were cryoprotected using the crystallization condition plus 20% glycerol and flash frozen in liquid nitrogen previously to data collection experiment. Data were collected at 100 K on ID23-2 beamline of the European Synchrotron Radiation Facility (ESRF, Grenoble, France).

For *Bs*UDG crystallization, an ammonium sulphate pellet of the protein was resuspended in buffer A to a final concentration of 14.7 mg/ml and dialyzed against the same buffer. Crystallization trials were performed as before using two commercial screens, (PACT Suite and JCSG+, both from Qiagen). Tiny rod crystals appeared after 3 weeks in a condition containing 0.2 M NaF and 20% PEG 3350 at 18°C. Further optimization of this condition did not progress; therefore, these crystals were cryoprotected using the crystallization condition supplemented with 20% glycerol. A data set was collected at 100 K in the ID29 beamline of the ESRF.

Structure determination, refinement and analysis of *Bs*UDG-p56 complex and *Bs*UDG

*Bs*UDG-p56 crystal data processing and reduction were done with iMosflm (26) and Scala (27) from the CCP4 suite (28). Crystals belong to the space group $P2_12_12_1$, with unit cell parameters a = 53.77, b = 66.27, c = 102.28 Å. There is one molecule of *Bs*UDG and a dimer of p56 in the asymmetric unit. The crystal structure of the *Bs*UDG-p56 complex was determined by Molecular Replacement using MolRep (29) and UDG from *E. coli* as initial search model (Protein Data Bank (PDB) code 2eug, 53% sequence identity). The model obtained was mutated to *Bs*UDG sequence and refined with the rigid body protocol of Refmac5 (30). This procedure led to a model that provided initial phases of good quality to trace the p56 dimer in the electron density maps by Buccaneer (31). Cycles of manual building using Coot (32) and refinement using Refmac5 (30) were performed. The electron density map showed double conformation for several protein residues as well as density for a chloride ion and a glycerol molecule, which have been modelled. Statistics for data processing and structure solution and refinement are shown in Supplementary Table S2. Final model includes residues 3–225 of *Bs*UDG, residues 8–55 of p56 subunit A and residues 4–56 of p56 subunit B.

*Bs*UDG crystal data set was processed with X-Ray Detector Software (33), and data were reduced with Scala (27) from the CCP4 suite (28). Crystals belong to space I222 group, with unit cell parameters a = 65.30, b = 79.79, c = 97.36 Å and one protein molecule in the asymmetric unit. The structure was determined by Molecular Replacement using Phaser (34). The high-resolution structure of UDG from *B. subtilis* as seen in complex with p56 protein, was used as initial search model. Structure refinement was performed alternating cycles of model rebuilding with Coot (32) and refinement with Refmac5 (30). Statistics for data reduction and refinement are shown in Supplementary Table S2.

Coordinates validation has been performed with PROCHECK (35). Phe78, a residue involved in conforming the uracil-binding pocket, violates the Ramachandran distribution, a characteristic conserved in other UDG known structures. The accessible surface areas and interface analysis have been computed with PISA server (36), and the pictures have been made with PyMOL (37).

Modelling of *Bs*UDG-DNA complex

*Hs*UDG was described to exist in open (free) and closed (complexed with DNA) conformations. Therefore, to

model the *Bs*UDG-DNA complex, *Bs*UDG coordinates were separated in two domains [domain I (8–79; 120–154) and domain II (80–111; 159–225], as their equivalent domains in several UDG isoforms are reported to close on DNA binding (38). Each domain was superposed to *Hs*UDG (50% sequence identity), as obtained in complex with DNA (pdb code 1emh). The linkers between domains I and II, and between segments in domain II were regularized using Coot (32). The model obtained presented a few close contacts at the protein–protein interface. Then, to improve the model quality, it was subjected to structure idealization using Refmac (30). The final model presents 90% residues inside the most favoured regions of the Ramachandran distribution, and none close contacts, as analysed by PROCHECK (35), conserving the general features of the *Hs*UDG–DNA complex.

The structural data of proteins from this publication have been submitted to the Protein Data Bank and assigned the identifier accession codes 3zoq and 3zor for *Bs*UDG-p56 and *Bs*UDG structures, respectively.

RESULTS

Structure of *Bs*UDG-p56 complex

We have crystallized the complex formed between the phage ϕ29 protein p56 and the UDG enzyme from *B. subtilis* (*Bs*UDG) at 1.45 Å resolution (Figure 1A and Supplementary Table S2; see amino acid sequence in Supplementary Figure S1). The folding of UDG from *B. subtilis* is conserved with respect to UDGs from other organisms previously crystallized (11–16). UDG is a classic α/β protein that consists of four parallel β-strands forming a β-sheet flanked by several α-helices on both sides (Figure 1B). The active site is situated in a wide cleft spanning the C-terminus of the β-sheet whose inner strands (β1 and β3) are noticeably opened. This cavity is surrounded by several conserved loops named LI, LII, LIII, LIV and LV (Figure 1B).

Protein p56, as seen in the *Bs*UDG-p56 complex, forms a tight dimer (subunits A and B; rmsd = 0.423 Å). This result is in agreement with previous analytical centrifugation studies showing that p56 formed dimers under

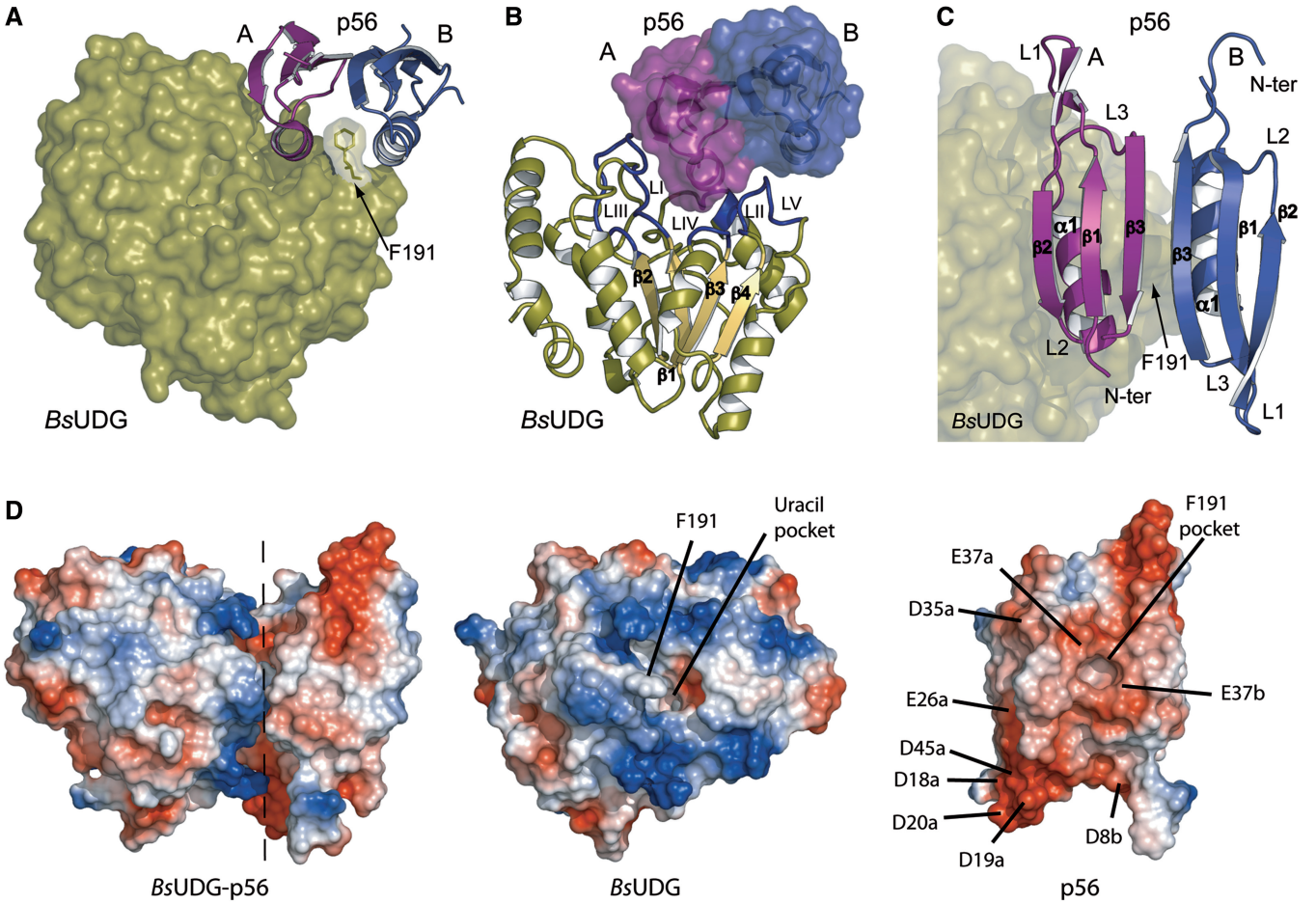

Figure 1. Structure of *Bs*UDG-p56 crystal complex. (**A**) Protein p56 subunits are represented as magenta (subunit A) and blue (subunit B) cartoons, whereas *Bs*UDG is shown as olive surface, transparent on residue Phe191. This code of colours is kept throughout the manuscript. (**B**) Protein p56 is shown as surface and *Bs*UDG as cartoon, showing the loops LI-LV in blue colour. (**C**) Zoom of figure (A), in a perpendicular orientation. (**D**) Electrostatic surface of: *Bs*UDG-p56 complex (left), the interface being highlighted by a dashed line; *Bs*UDG (middle), showing the basic character of the p56 binding surface, and p56 (right), showing the acidic character of the *Bs*UDG binding surface in which the position of acidic residues is shown.

physiological conditions (24) and with Isothermal Titration Calorimetry analysis revealing that p56 complexed with *Bs*UDG in 2:1 stoichiometry (39). Each monomer consists of three antiparallel β-strands and one α-helix, where β1 is the inner strand, and α1 is linking the outer strands β2 and β3 (Figure 1C). These secondary structural elements are connected by several loops named L1, L2 and L3. The elements from p56 subunits A and B are named, respectively, with the suffixes a and b therein, to distinguish them from each other. The two p56 subunits, related by a pseudo 2-fold axis, form a dimer by symmetric interactions through α1a-α1b and β3a-β3b producing an extended β-sheet with six antiparallel βstrands (Figure 1C).

The p56 regions involved in *Bs*UDG recognition exhibit a negative electrostatic potential, compatible with the highly basic character of the DNA-binding loops at the enzyme interface (Figure 1D). Interestingly, p56 subunits bind differently to *Bs*UDG generating a non-symmetrical complex. Subunit A of protein p56 produces a much more extensive interaction with *Bs*UDG contributing with 1510 Å^2 to the total buried surface of the complex calculated in 1815 Å^2 (Supplementary Table S3, see Figure 2A). The recognition between *Bs*UDG and p56 involves multiple polar (summarized in Supplementary Table S4) and hydrophobic interactions (Figure 2A). Our analysis revealed that the main feature of the complex formation is the recognition of Phe191 from *Bs*UDG by p56. Both p56 subunits use their α1 as clamps to completely enclose Phe191 in a hydrophobic pocket located at the dimerization interface (Figure 2B). Residues Glu37, Phe36 and Tyr40 from both p56 α1 outline this pocket, representing the only symmetrical interaction of the p56 dimer with *Bs*UDG. In addition, Glu37 and Tyr40 play an essential role in p56 dimerization (39), as the two hydrogen bonds between Glu37a-Tyr40b and Glu37b-Tyr40a are crucial in maintaining its architecture.

The interface between p56 subunit B (p56B) and *Bs*UDG is formed exclusively by the hydrophobic contact through Phe191, plus the polar interaction between Ser34b and Glu37b from p56 and Arg194 from *Bs*UDG (Figure 2A). However, as aforementioned, subunit A of p56 (p56A) produces a much more extended interaction with *Bs*UDG. Besides the interaction mediated by the *Bs*UDG Phe191, other hydrophobic contacts were observed along the whole *Bs*UDG-p56A interface (Figure 2C) like those produced between Tyr67, His68 and His187 from *Bs*UDG and Asn42a, His28a and Val38a from p56A. Moreover, several hydrophobic contacts with p56A involve *Bs*UDG prolines, in particular Pro87 and Pro88 from the proline rich loop, and Pro190, which interact with the region Y40a-N42a of p56.

Nevertheless, there are also multiple polar interactions implicated in the complex formation (Supplementary Table S4). In all, 11 of 19 total polar interactions found in the complex involve residues from p56 α1, either α1a (Figure 2D) or α1b (Figure 2A), but other p56 regions also present relevant interactions (Figure 2E). Among the polar interactions, those produced between *Bs*UDG LII (Pro88 and Ser89) and p56 α1 (Gly41a) are particularly significant, as they involve backbone links (Figure 2D). In addition, four basic residues of *Bs*UDG, Lys85, Lys137, Arg166 and Arg194 form several ion pairs that are implicated in the complex formation (Figure 2F). Specifically, Arg166 and Arg194 interact with Glu37 in both α1a and α1b helices, respectively, whereas Lys185 forms an ion net with two acidic residues of p56, Asp19a and Asp45a, both of which are outside of α1 (Figure 2A and 2D–F). In addition to the direct polar links, multiple interactions ($17 < 3.2$ Å) through water molecules are formed within the *Bs*UDG-p56 interface, increasing the strength of the interaction produced between both proteins.

In summary, the main feature of the *Bs*UDG-p56 complex is the recognition of *Bs*UDG Phe191 within a hydrophobic pocket at the p56 dimer interface. *Bs*UDG residues from LV and the remaining DNA-binding loops seal this interaction through polar and hydrophobic interactions established mainly with p56 α1a.

*Bs*UDG and p56 structural changes on complex formation

To characterize the structural changes that occur in *Bs*UDG on p56 binding, we have crystallized UDG from *B. subtilis* and solved its structure at 2.9 Å resolution, using the high resolution *Bs*UDG model as found in the *Bs*UDG-p56 complex. The final refined *Bs*UDG model showed unambiguous density for the whole molecule. The superposition of the *Bs*UDG structure free or in complex with p56 (rmsd = 0.382 Å) reveals variations in two of the five *Bs*UDG loops mentioned previously (Figure 3A). The most significant change occurs at LV, which contains the Phe191, a residue that is completely buried in a hydrophobic pocket in the complex while it is exposed to the solvent and partially disordered in the absence of p56. Loop II, containing several prolines, also displays an appreciable variation reflecting its intrinsic flexibility. The role of proline residues in protein–protein complex formation has been described in previous studies (40,41). Despite these modifications, the global *Bs*UDG structure does not seem to change significantly on p56 binding, revealing that p56 inhibits the enzyme in its free state conformation.

Similarly, the p56 fold and dimerization features found in the *Bs*UDG-p56 crystal complex are essentially the same as those described for free p56 recently solved by nuclear magnetic resonance (NMR) (39) (Figure 3B) (rmsd = 1.16 Å with respect to the top NMR model: pdb code 2le2).

In summary, the interaction surface is nearly preformed in both proteins, suggesting that no additional energetic cost is required to form the complex, and providing a possible explanation for its tight binding (24).

Effect of *Bs*UDG mutations on enzymatic activity and DNA-binding capacity

Several data suggest that p56 mimics the interactions between *Bs*UDG and the phosphate backbone of the DNA. To confirm this possibility, we obtained a structural model of the *Bs*UDG–DNA complex (Figure 4) using the

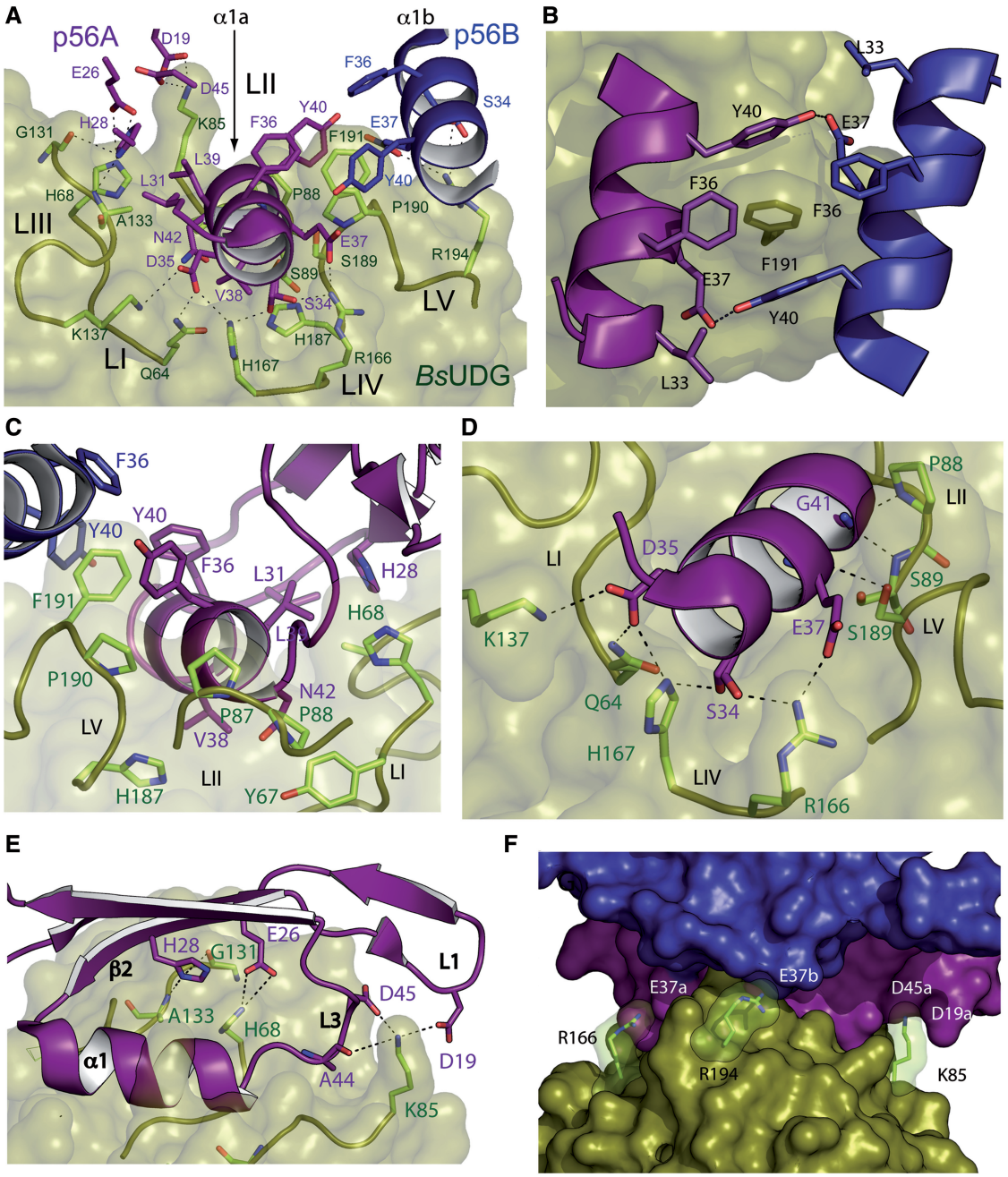

Figure 2. *Bs*UDG-p56 interaction. (**A**) p56-*Bs*UDG interface showing all the interactions produced in the complex. *Bs*UDG is shown as transparent surface, highlighting the interacting loops, while p56 α-helices are shown as cartoons. The residues making interactions are shown as sticks. (**B**) *Bs*UDG Phe191 recognition by A and B subunits of p56. Phe 191 binds into a symmetrical hydrophobic pocket situated in the p56 dimer interface. (**C**) Scheme of *Bs*UDG-p56 hydrophobic interactions. (**D**) Scheme of polar interactions formed between *Bs*UDG and α1 of p56A. (**E**) Scheme of polar interactions produced between *Bs*UDG and p56A elements outside α1. (**F**) Surface of the *Bs*UDG-p56 complex, showing three basic residues of *Bs*UDG that make ionic pairs with p56 acidic residues. Residue Lys137 cannot be shown in this orientation.

*Hs*UDG–DNA complex as a template (pdb code: 1emh, 50% sequence identity). This model allowed us to propose the most relevant *Bs*UDG residues involved in DNA binding and to compare them with those implicated in p56 binding. To gain a further insight into the biochemical basis of the complex formation, we performed a site-directed mutational analysis on *Bs*UDG.

*Bs*UDG mutants containing the single mutation Q64A, H68A, K85S, R166S, H187A, S189A, P190A, F191A or R194S, or the double mutation P87A/P88A were expressed and purified to apparent homogeneity, and their enzymatic activity as well as their DNA-binding ability were determined. The UDG activity was significantly decreased in all the mutants (Figure 5A and Supplementary Figure S2), although they can be classified into three groups according to their residual enzymatic capacity. The first group is formed by mutants K85S and R166S with a uracil excision capacity ~3-fold

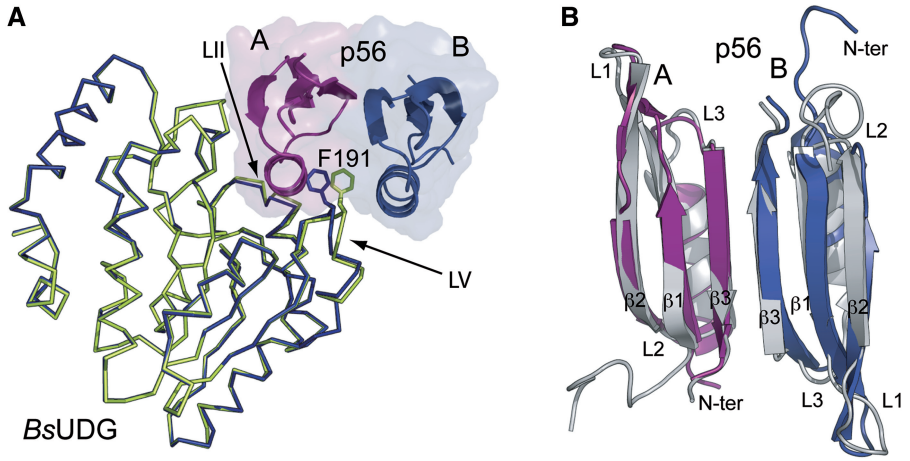

Figure 3. Changes produced in *Bs*UDG and p56 upon complex formation. (**A**) Superposition of *Bs*UDG free and in complex with p56. *Bs*UDG is represented as Cα trace in colour blue (free) and green (p56-bound). Protein p56 is shown as cartoons surrounded by a transparent surface. The most variable regions are marked by arrows, in which Phe191 is shown as sticks. (**B**) Superposition of p56 structures from *Bs*UDG-p56 crystal complex and from NMR structure (coloured grey).

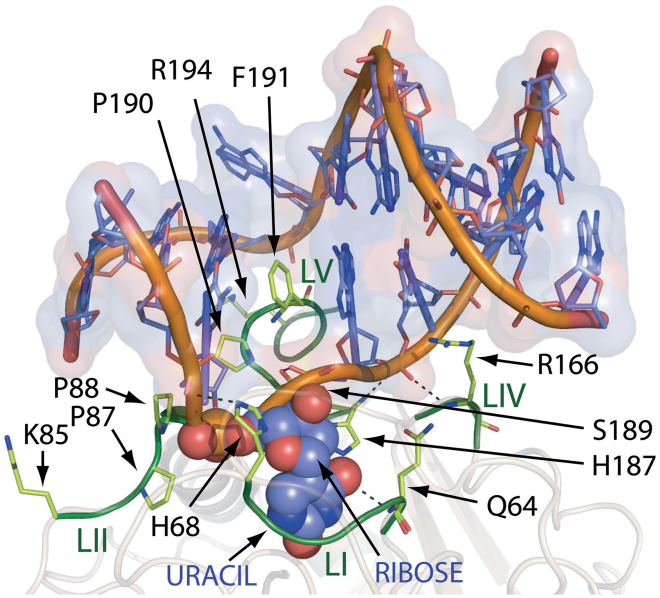

Figure 4. Model of *Bs*UDG-DNA interacting surface. DNA backbone is shown as orange cartoons, whereas the bases are shown as blue sticks. The flipped uracil is shown as spheres. Oxygen and nitrogen atoms are coloured in red and blue, respectively. The binding interface of *Bs*UDG is shown as transparent cartoons highlighting the interacting loops in green colours. The residues mutated for this work are shown as green sticks in the above loops.

reduced with respect to wild-type *Bs*UDG (Figure 5A). These residues are not conserved among UDGs from different organisms and are located at the boundaries of the *Bs*UDG-p56 complex (Figure 2F). Curiously, electrophoretic mobility shift assays showed that K85S and R166S were the mutants most affected in DNA binding (Figure 5B and Supplementary Figure S3), as they retained only <40% of the wild-type protein binding. Previous studies have proposed that basic residues on *Hs*UDG surface have a role in attracting DNA (11). Thus, the moderate reduction in enzymatic activity in these mutants could be a consequence of the decrease in DNA binding. The second group of mutants, formed by H68A, P87A/P88A, S189A and R194S, have an activity decrease ~6–10-fold when compared with wild-type *Bs*UDG. With the exception of Arg194, these residues are located at the entrance of the active site to ensure a proper DNA positioning, in particular that of uracil. The equivalent residues of His68 and Ser189 in *Hs*UDG (His148 and Ser270, respectively) bind to the DNA phosphates, whereas conserved prolines 87 and 88 (Pro167 and Pro168 in *Hs*UDG) pack against DNA to hold the 5′-phosphate from the uracil-containing nucleotide. Although far from the active site, the equivalent residue of Arg194 in *Hs*UDG (Tyr275) stacks with DNA bases making a hydrogen bonding interaction (42). Despite the fact that *Bs*UDG activity was clearly affected, the mutation of these residues did not produce a substantial decrease in DNA binding (Figure 5B and Supplementary Figure S3) probably because the specific recognition between *Bs*UDG and DNA involves multiple interactions through several residues that contribute to complex formation. The third group of mutants (Q64A, H187A, P190A and F191A) was severely impaired, as they showed no detectable UDG activity (except in the case of P190A, which retained a 4% activity) even when the amount of mutant proteins was increased 10-fold compared with that of the wild-type. Other studies demonstrated that *Hs*UDG residues Gln144 and His268 (equivalent to *Bs*UDG Gln64 and His187) interact with uracil oxygens through its main and side chains, respectively. In addition, Gln144 is necessary to orient the uracil properly into the active site through its side chain, whereas His268, a catalytic residue proposed to be involved in the uracil excision process (11), coordinates a DNA phosphate through its main chain. Phe191 is also essential for *Bs*UDG function and is equivalent to Leu272 of *Hs*UDG. The side chain of Leu272 is inserted into the DNA minor groove protruding into the DNA base stack to replace the flipped-out uracil nucleotide. The

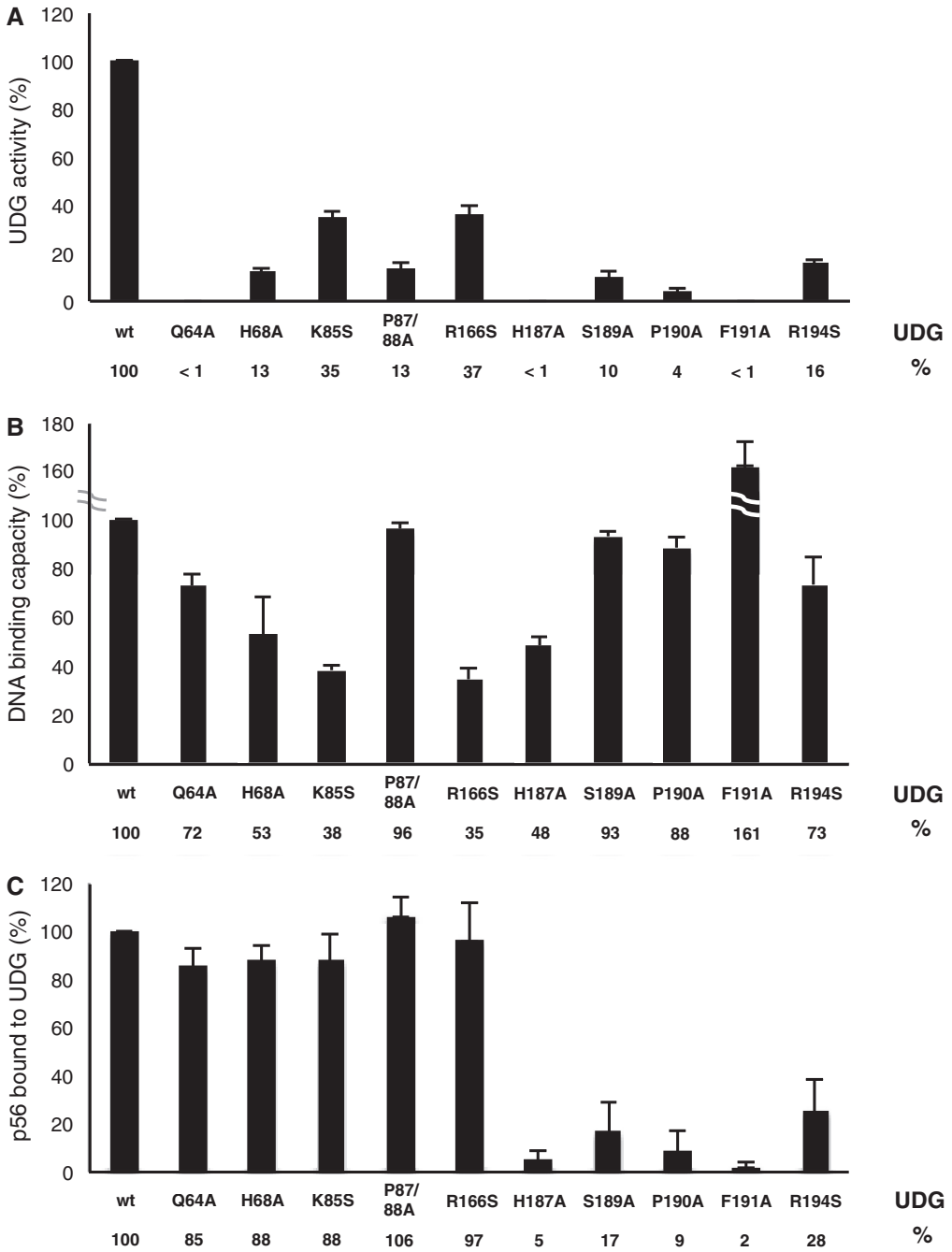

Figure 5. Influence of the mutated residues of *Bs*UDG on the enzymatic activity, DNA-binding capacity and complex formation with p56. (**A**) Enzymatic activity of *Bs*UDG mutants compared with wild-type *Bs*UDG. The 34 bp 5′-^{32}P-labelled ssDNA substrate containing a single uracil residue at position 16 was incubated with the different proteins. Bands corresponding to DNA substrate and DNA product were monitored by autoradiography. (**B**) The DNA-binding ability of the GST-*Bs*UDG wild-type and the mutants was examined by gel retardation assays. Proteins were incubated in binding buffer with 5′-^{32}P-labelled dsDNA U:G substrate, and the samples were analysed by non-denaturing polyacrylamide gel electrophoresis. Bands were detected by autoradiography. (**C**) Role of the mutated residues of GST-*Bs*UDG in the interaction with protein p56 as determined by pull-down assay. After incubation, the quantity of p56 bound to GST-*Bs*UDG proteins was determined by Tris–Tricine polyacrylamide gel electrophoresis and visualized using Coomassie blue staining. Bands were quantified by densitometry. The percentage of UDG activity (A), DNA-binding capacity (B) or p56-binding ability (C) for the mutants was calculated with respect to the wild-type *Bs*UDG. Each bar represents the mean value for three independent experiments with error bars indicating the standard deviations.

replacement of Leu272 with alanine (L272A) practically abolished UDG activity (14), as it occurs in *Bs*UDG F191A mutant. The contiguous *Hs*UDG Pro271 (equivalent to *Bs*UDG Pro190) packs with a base preceding the flipped nucleotide, ensuring a correct DNA location. All these mutants conserved at least ~50% of the wild-type ability to bind DNA, and therefore, it seems unlikely that the selected mutations have a significant influence on the structural stability of these proteins. In conclusion, all the *Bs*UDG residues mutated from the *Bs*UDG-p56 interface

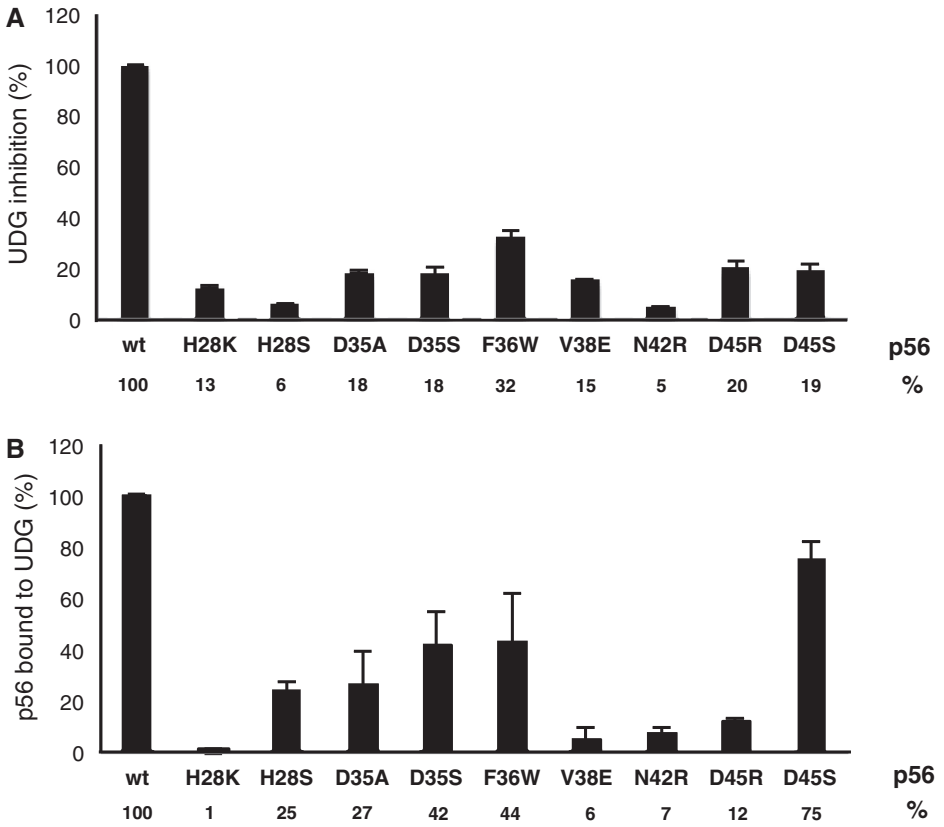

Figure 6. Role of the mutated residues of protein p56 in UDG inhibition and binding capacity to *Bs*UDG. (**A**) Ability of the p56 mutants to inhibit the UDG activity. p56 mutants and wild-type *Bs*UDG were incubated and then the 34 bp 5′-^{32}P-labelled ssDNA substrate containing a single uracil residue at position 16 was added to the reaction. Samples were subjected to electrophoresis in polyacrylamide gels, and bands were detected by autoradiography. (**B**) Influence of the mutated residues of p56 in the interaction with protein GST-*Bs*UDG as determined by pull-down assay. Following incubation the quantity of p56 bound to GST-*Bs*UDG proteins was examined by Tris–Tricine polyacrylamide gel electrophoresis and visualized using Coomassie blue staining. Bands were quantified by densitometry. The percentage of UDG inhibition (A) or UDG-binding capacity (B) for the mutants was calculated relative to the wild-type p56. Bar charts represent the average values for three independent experiments with standard deviations indicated.

are important for enzyme activity. In particular, residues located at the uracil-binding pocket entrance (Gln64 and His187) and Phe191, a residue that protrudes into the DNA minor groove, are essential, in agreement with the structural and functional findings of homologue residues in other UDGs (19).

Amino acids involved in the binding of p56 to *Bs*UDG and its inhibition

To determine the ability of *Bs*UDG mutants to form a complex with p56, we performed GST pull-down experiments. The amount of each GST-*Bs*UDG mutant in complex with p56 was measured following incubation under standard binding conditions as described in 'Materials and Methods' section. As shown in Figure 5C and Supplementary Figure S4, the stability of the *Bs*UDG-p56 complex varies considerably among the mutants. Interestingly, the most affected *Bs*UDG mutants (H187A, S189A, P190A, F191A and R194S) contained point mutations in residues located at loop V. This is consistent with the observation that the main interaction between *Bs*UDG and p56 is produced through Phe191 binding, a residue located in the middle of LV. Consequently, F191A almost prevents p56 binding. Other LV mutants, such as H187A and P190A, also show a reduced p56 binding capacity. These residues are involved in hydrophobic contacts with α1a of p56 (residues Val38a and Tyr40a, respectively), indicating the importance of these hydrophobic interactions in the complex formation. Mutants S189A and R194S still showed a significant reduction in p56 binding (84 and 75%, respectively). Both mutated residues interact with Glu37 from p56 subunit A and B, respectively (Figure 2). Although the Ser189 interaction is produced through its main chain, the intramolecular contact produced between His187 and Ser189 side chains can explain the large effect of the S189A mutation, as this residue seems to participate in a specific UDG surface configuration and/or His187 positioning.

In a complementary approach, we performed a mutational analysis of p56 to investigate the role of the selected residues in both complex stability and inhibitory capacity of the UDG activity. All the p56 mutants tested (H28K, H28S, D35A, D35S, F36W, V38E, N42R, D45R and D45S) were noticeably less efficient than the wild-type p56 at preventing the cleavage of uracil-containing substrate by *Bs*UDG (Figure 6A and Supplementary Figure

S5). Specifically, these substitutions led to ~80% reduction in the inhibitory activity with respect to the wild-type protein with the exception of mutant F36W that produced a 3-fold decrease in inhibition. However, Phe36 is expected to be essential in both p56 dimerization and Phe191 binding site formation (Figure 2B), suggesting that although substantial distortion may be introduced by the substitution, the tryptophan side chain probably keeps the main features of Phe36. The mutants on the acidic residues Asp35 and Asp45 exhibited a 5-fold inhibition decrease. Both residues form hydrogen bonds with *Bs*UDG. Asp35a is able to form links to Asn64, His167 and Lys137 from *Bs*UDG (Figure 2D), whereas Asp45a interacts with Lys85 and also makes an intramolecular contact with Tyr24a (Figure 2E) contributing to the shaping of the interface. D45R and D45S showed similar activity, thus suggesting that arginines can be easily accommodated on the surface without causing any deleterious effects. The mutation of the hydrophobic residue Val38 presents a similar effect to the aforementioned acidic residues. Val38a stacks to *Bs*UDG His187, and, unusually, it points to a polar cavity containing several water molecules (Supplementary Figure S6A). However, the V38E change probably affects its interaction with His187 or the conformation of neighbour residues, decreasing the binding capacity. The most affected mutants are H28K, H28S and N42R. His28a makes both polar (Ala133) and hydrophobic (His68) contacts with *Bs*UDG (Figure 2E and C). In addition, His28 is involved in the configuration of the interaction surface through intramolecular interactions with Leu39 and Glu26, residues that also interact with *Bs*UDG (Supplementary Figure S6B). Finally, Asn42a is the only p56 residue that points towards the uracil-binding pocket, though it does not protrude deeply into the cavity. The lack of inhibition in the N42R mutant could be explained if the charge introduced is not well accommodated in a partially hydrophobic pocket, despite the fact that there seems to be enough room to accommodate a large residue. In addition, Asn42a is involved in several water-mediated interactions through its main and side chain, and its elimination could also affect binding.

The significant decreased capacity of p56 mutants to inhibit *Bs*UDG activity seems to reflect a reduced ability to bind to *Bs*UDG (Figure 6B and Supplementary Figure S7), suggesting a role for these residues in the formation of the complex. Interestingly, p56 mutants showed relevant differences in their binding to *Bs*UDG despite their fairly similar inhibitory capacity. Mutants can be classified into three groups according to their ability to form stable complexes with *Bs*UDG. The first group includes the most affected mutants (H28K, V38E, N42R and D45R) that only form <10% of the complexes relative to the wild-type protein. The second one is constituted by mutants H28S, D35A, D35S and F36W that retain ~25–45% of the wild-type p56 binding to *Bs*UDG, whereas D45S, the only member of the third group, maintains ~75% of the wild-type complex formation capacity. The high specificity of the p56 inhibition is based on its ability to bind to UDG more tightly than the DNA substrate (24). The moderate *Bs*UDG binding capacity displayed by some p56 mutants does not seem to be enough to dissociate DNA from *Bs*UDG, as they were not capable of producing any significant inhibition of UDG activity. Therefore, these results may suggest that p56 mutants, even those less severely affected in *Bs*UDG binding, are not able to compete with DNA for binding to UDG as efficiently as the wild-type p56.

DISCUSSION

Uracil DNA-glycosylases are highly conserved enzymes that play a crucial role in maintaining genomic integrity by efficiently removing uracil from DNA (43). Nevertheless, two bacteriophage proteins, named Ugi and p56, have been characterized as natural inhibitors of UDGs from a variety of different organisms, including their host *B. subtilis*. Intriguingly, both inhibitors act as DNA mimic proteins that might have evolved to inactivate *Bs*UDG activity, representing a viral defence mechanism to overcome the host uracil excision repair pathway. Specifically, the phage PBS1/PBS2-encoded Ugi is essential for virus viability by impeding the excision of uracil residues from the unusual uracil-containing DNA of the phage (18). In addition, we have recently reported that phage ϕ29 protein p56 prevents viral DNA replication impairment caused by the removal of any uracil erroneously incorporated into the replicative intermediates (24,44).

In the present study, we have solved the crystal structures of both uncomplexed and complexed *Bs*UDG with phage ϕ29 protein p56 to a resolution of 2.9 and 1.45 Å, respectively, to analyse the structural basis for enzyme inhibition. The most significant interactions in the formation of *Bs*UDG-p56 complex were identified and compared with those in the structural model of *Bs*UDG-DNA. Moreover, we have taken advantage of the fact that UDG is the only enzyme known to be inhibited by two different DNA mimic proteins to compare both p56 and Ugi interactions with UDG (Figure 7A). This approach allowed us to identify the structural and biochemical properties shared by both inhibitors that enable them to block the DNA-binding domain of *Bs*UDG.

Previous studies have determined that the mechanism for uracil repairing by UDGs in diverse organisms included the displacement of the uracil lesion to an extrahelical position and the insertion into the DNA base stack of the side chain of a protruding leucine (Leu272 in *Hs*UDG, Leu191 in *Ec*UDG) located at UDG loop V (14,38,45). Interestingly, our analysis revealed that the main feature of the *Bs*UDG-p56 complex formation is the recognition of the Phe191 residue from *Bs*UDG by a specific hydrophobic pocket at the p56 dimer interface (Figure 7B). Residues located at the DNA-binding loops of *Bs*UDG, particularly those in loop V, are responsible for sealing this interaction through polar and hydrophobic interactions (Figure 2A). Other studies have described that Ugi also envelops a protruding leucine residue from human and *E. coli* UDG (19,46,47). Functional characterization of specific *Bs*UDG mutants showed that the most important residues for p56 binding (His187, Ser189,

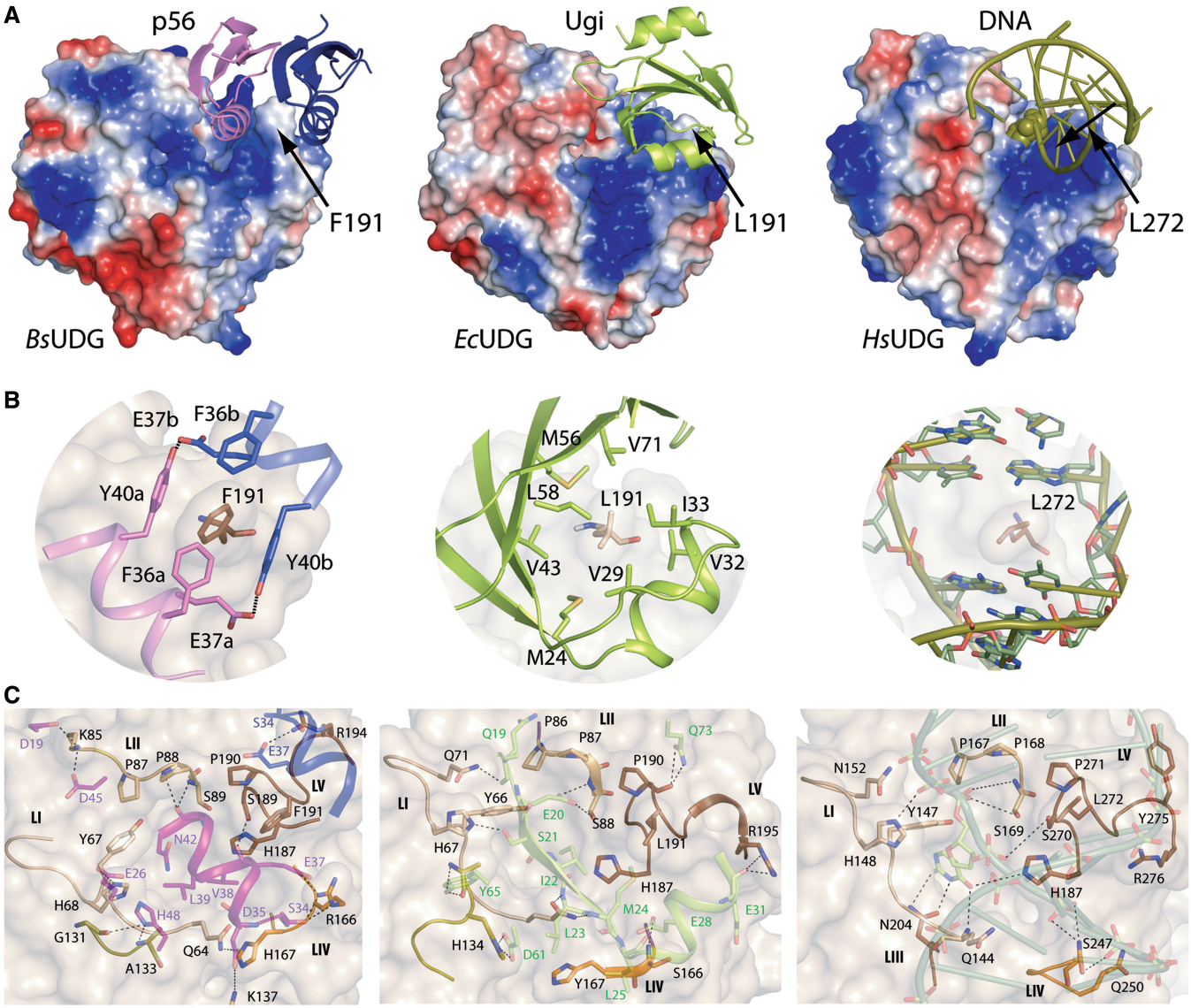

Figure 7. Comparison of p56, Ugi and DNA-binding mode in UDG proteins. (**A**) Structure of *Bs*UDG-p56 (left), *Ec*UDG-Ugi (middle) (pdb code: 1ugh) and *Hs*UDG-DNA (right) (pdb code: 1 emh) complexes. *Hs*UDG-Ugi complex keeps the general features of *Ec*UDG-Ugi. In all cases, the UDG protein is shown as electrostatic surface, and the inhibitory proteins or DNA are shown as cartoons in magenta/blue, green and olive colours, respectively. (**B**) Zoom showing the Phe191 (left) /Leu191 (middle) /Leu272 (right) binding pockets. UDGs are shown as cream transparent surface, and p56/Ugi/DNA elements as cartoons with the essential residues highlighted as sticks. (**C**) Detail showing the five UDG loops (LI–LV) involved in binding in different brown shades. Proteins p56, Ugi or DNA important elements for UDG binding are shown as transparent cartoons. Essential residues for complex formation are shown as sticks and the polar interactions marked as dashed lines.

Pro190 and Phe191) were also essential for the enzymatic activity (Figures 5A and 6C). The equivalent residues in *Hs*UDG are also implicated in the uracil excision process. In particular, it is thought that the insertion of the Leu272 side chain (Phe191 in *Bs*UDG) into the DNA minor groove to replace the flipped-out uracil allows His268 (His187 in *Bs*UDG) to move to a catalytic location (11). Ser270 (Ser189 in *Bs*UDG) and Pro271 (Pro190 in *Bs*UDG) mediate in this process by ensuring a correct uracil positioning through interactions with the preceding nucleotide base and subsequent nucleotide phosphate, respectively.

Unexpectedly, the substitutions of the selected residues clustered in loop V notably reduce the capacity of binding p56 without significantly altering the DNA-binding ability of the enzyme. The biochemical properties of the interaction between these critical *Bs*UDG residues and their counterparts in p56 provide a plausible explanation for the fact that UDG fits its inhibitor more tightly than the DNA substrate, as protein p56 is able to dissociate preformed UDG–DNA complexes (24). Positively charged *Bs*UDG amino acids, such as Lys85 from loop II and Arg166 from loop IV, seemed to be more relevant than loop V residues for binding to the DNA substrate (Figure 5B).

Our studies suggest that the two helical regions of the p56 dimer (α1a and α1b) fit into the cavities where both DNA strands are proposed to bind to *Bs*UDG (Figure 7C).

Ugi reaches the equivalent region in *Hs*UDG and *Ec*UDG through its β1 strand, a structure with an uncommon shape that resembles the bent twist of the DNA in complex with UDG (19,46). However, neither p56 nor Ugi interact directly with residues of the UDG uracil-binding pocket (19). Based on these results, we propose that protein p56 might target the conserved mechanisms for uracil flipping and amino acid intercalation into the base stack. Thus, p56 probably mimics just a short DNA fragment, as UDGs only seem to interact with the extrahelical uracil and the phosphate backbone flanking the flipped base during the catalytic activity. Previously, it has been suggested that protein Ugi could target the same step in the enzymatic reaction for UDG inhibition (46). Therefore, both inhibitors possibly prevent non-specific cross-reactive events with other DNA-binding enzymes by targeting the specific mode of action used mainly by UDGs.

The binding to uracil-containing DNA induces a conformational change in UDGs from an 'open' to a 'closed' state that generates the enzyme active site (42). This movement in *Hs*UDG is described as a 10° rotation (38). On the contrary, only discrete and local structural changes occur in *Bs*UDG and p56/Ugi proteins on complex formation, suggesting that the shape and electrostatic complementarity is pre-existing [Figure 3; (46)]. This is consistent with the tighter binding of p56/Ugi to UDG than DNA. As expected for DNA mimic proteins, p56 and Ugi regions implicated in UDG recognition exhibit a negative electrostatic potential, compatible with the positively charged surface of the UDG interface [Figure 1D; (11)]. In agreement with these structural data, mutagenesis analysis of p56 acidic residues located in the complex interface have confirmed the key role of Asp35 and Asp45 (this work), and Glu26, Glu37 and Asp18/Asp19/Asp20 (39) in complex formation. Previous studies have also shown that the mutation of two DNA phosphate-mimicking Ugi carboxylate groups, Glu20 and Glu28, impeded the formation of a stable complex with UDG, and thus, the enzyme activity was not completely inhibited (48).

Our results provide insights into the different behaviour of the two inhibitors, as *Ec*UDG-Ugi and *Hs*UDG-Ugi complexes generate a larger and more specific interface in addition to a more favourable binding energy on their formation than the *Bs*UDG-p56 complex (Supplementary Table S5). Moreover, the recognition carried out by Ugi of the protruding UDG residue seemed to occur over a larger binding surface compared with that in p56. Remarkably, a comparison between the experimental enzyme-inhibitor structures with several theoretical models of p56 and Ugi binding to the same UDG isoform, suggested that higher polar and electrostatic complementarity was found in UDG-Ugi complexes (Supplementary Methods and Supplementary Table S5). Both inhibitors show an expected enhanced fit to UDG from their natural host, *B. subtilis*, probably as a result of an increase in the number of polar interactions and especially, ion-pairs contacts (Supplementary Table S5).

Altogether, our results provide new insights in understanding the basis for the binding specificity of DNA mimic proteins, and they could also contribute to the design of selective UDG inhibitors with different properties.

ACCESSION NUMBERS

PDB 3zoq and 3zor.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–5, Supplementary Figures 1–7, Supplementary Methods and Supplementary References [25,35,36,49].

ACKNOWLEDGEMENTS

The authors thank the European Synchrotron Radiation Facility (Grenoble, France) for providing beam time and ID23-2 and ID29 staff for assistance during data collection.

FUNDING

Spanish Ministry of Economy and Competitiveness [BFU2011-23645 to M.S., BFU2011-24982 to B.G.]; CONSOLIDER-INGENIO from the Spanish Ministry of Science and Innovation [CSD2007-00015 to M.S.]; Fundación Mutua Madrileña (to M.S.); Madrid Autonomous Community [S2009MAT-1507 to M.S.]; Fundación Ramón Areces to the Centro de Biología Molecular 'Severo Ochoa' Institutional Grant; Spanish Ministry of Science and Innovation [BIO2010-20508-C04-03 to J.S.-A.]; Formación de Profesorado Universitario Fellowship [AP2008-00916 to J.I.B.-S.], Spanish Ministry of Education and Culture. L.M. was holder of a CONSOLIDER-INGENIO contract. Funding for open access charge: CONSOLIDER-INGENIO [CSD2007-00015].

Conflict of interest statement. None declared.

REFERENCES

1. Hakem,R. (2008) DNA-damage repair; the good, the bad, and the ugly. *EMBO J.*, **27**, 589–605.
2. Branzei,D. and Foiani,M. (2008) Regulation of DNA repair throughout the cell cycle. *Nat. Rev. Mol. Cell Biol.*, **9**, 297–308.
3. Aguilera,A. and Gómez-Gónzalez,B. (2008) Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.*, **9**, 204–217.
4. Friedberg,E.C. (1995) Out of the shadows and into the light: the emergence of DNA repair. *Trends Biochem. Sci.*, **20**, 381.
5. Andersen,S., Heine,T., Sneve,R., Konig,I., Krokan,H.E., Epe,B. and Nilsen,H. (2005) Incorporation of dUMP into DNA is a major source of spontaneous DNA damage, while excision of uracil is not required for cytotoxicity of fluoropyrimidines in mouse embryonic fibroblasts. *Carcinogenesis*, **26**, 547–555.
6. Krokan,H.E., Drablos,F. and Slupphaug,G. (2002) Uracil in DNA-occurrence, consequences and repair. *Oncogene*, **21**, 8935–8948.
7. Shen,J.C., Rideout,W.M. 3rd and Jones,P.A. (1992) High frequency mutagenesis by a DNA methyltransferase. *Cell*, **71**, 1073–1080.
8. Lindahl,T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.

9. Lindahl,T., Ljungquist,S., Siegert,W., Nyberg,B. and Sperens,B. (1977) DNA N-glycosidases: properties of uracil-DNA glycosidase from *Escherichia coli*. *J. Biol. Chem.*, **252**, 3286–3294.
10. Pearl,L.H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.*, **460**, 165–181.
11. Mol,C.D., Arvai,A.S., Slupphaug,G., Kavli,B., Alseth,I., Krokan,H.E. and Tainer,J.A. (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**, 869–878.
12. Slupphaug,G., Mol,C.D., Kavli,B., Arvai,A.S., Krokan,H.E. and Tainer,J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87–92.
13. Barrett,T.E., Savva,R., Panayotou,G., Barlow,T., Brown,T., Jiricny,J. and Pearl,L.H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell*, **92**, 117–129.
14. Parikh,S.S., Mol,C.D., Slupphaug,G., Bharati,S., Krokan,H.E. and Tainer,J.A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.*, **17**, 5214–5226.
15. Xiao,G., Tordova,M., Jagadeesh,J., Drohat,A.C., Stivers,J.T. and Gilliland,G.L. (1999) Crystal structure of *Escherichia coli* uracil DNA glycosylase and its complexes with uracil and glycerol: structure and glycosylase mechanism revisited. *Proteins*, **35**, 13–24.
16. Kaushal,P.S., Talawar,R.K., Varshney,U. and Vijayan,M. (2010) Structure of uracil-DNA glycosylase from *Mycobacterium tuberculosis*: insights into interactions with ligands. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun*, **66**, 887–892.
17. Wong,I., Lundquist,A.J., Bernards,A.S. and Mosbaugh,D.W. (2002) Presteady-state analysis of a single catalytic turnover by *Escherichia coli* uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism. *J. Biol. Chem.*, **277**, 19424–19432.
18. Cone,R., Bonura,T. and Friedberg,E.C. (1980) Inhibitor of uracil-DNA glycosylase induced by bacteriophage PBS2. Purification and preliminary characterization. *J. Biol. Chem.*, **255**, 10354–10358.
19. Mol,C.D., Arvai,A.S., Sanderson,R.J., Slupphaug,G., Kavli,B., Krokan,H.E., Mosbaugh,D.W. and Tainer,J.A. (1995) Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. *Cell*, **82**, 701–708.
20. Ravishankar,R., Bidya Sagar,M., Roy,S., Purnapatre,K., Handa,P., Varshney,U. and Vijayan,M. (1998) X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (EcUDG) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG. *Nucleic Acids Res.*, **26**, 4880–4887.
21. Putnam,C.D. and Tainer,J.A. (2005) Protein mimicry of DNA and pathway regulation. *DNA Repair*, **4**, 1410–1420.
22. Acharya,N., Kumar,P. and Varshney,U. (2003) Complexes of the uracil-DNA glycosylase inhibitor protein, Ugi, with *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* uracil-DNA glycosylases. *Microbiology*, **149**, 1647–1658.
23. Serrano-Heras,G., Salas,M. and Bravo,A. (2006) A uracil-DNA glycosylase inhibitor encoded by a non-uracil containing viral DNA. *J. Biol. Chem.*, **281**, 7068–7074.
24. Serrano-Heras,G., Ruiz-Maso,J.A., del Solar,G., Espinosa,M., Bravo,A. and Salas,M. (2007) Protein p56 from the *Bacillus subtilis* phage ϕ29 inhibits DNA-binding ability of uracil-DNA glycosylase. *Nucleic Acids Res.*, **35**, 5393–5401.
25. Pérez-Lago,L., Serrano-Heras,G., Baños,B., Lázaro,J.M., Alcorlo,M., Villar,L. and Salas,M. (2011) Characterization of *Bacillus subtilis* uracil-DNA glycosylase and its inhibition by phage ϕ29 protein p56. *Mol. Microbiol.*, **80**, 1657–1666.
26. Battye,T.G., Kontogiannis,L., Johnson,O., Powell,H.R. and Leslie,A.G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 271–281.
27. Evans,P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 72–82.
28. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 235–242.
29. Vagin,A. and Teplyakov,A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 22–25.
30. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 355–367.
31. Cowtan,K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 1002–1011.
32. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
33. Kabsch,W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
34. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
35. Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
36. Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.
37. DeLano,W.L. (2002) *The PyMOL Molecular Graphics System*. DELano Scientific LLc, San Carlos, CA, USA.
38. Saikrishnan,K., Bidya Sagar,M., Ravishankar,R., Roy,S., Purnapatre,K., Handa,P., Varshney,U. and Vijayan,M. (2002) Domain closure and action of uracil DNA glycosylase (UDG): structures of new crystal forms containing the *Escherichia coli* enzyme and a comparative study of the known structures involving UDG. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1269–1276.
39. Asensio,J.L., Pérez-Lago,L., Lázaro,J.M., Gónzalez,C., Serrano-Heras,G. and Salas,M. (2011) Novel dimeric structure of phage ϕ29-encoded protein p56: insights into uracil-DNA glycosylase inhibition. *Nucleic Acids Res.*, **39**, 9779–9788.
40. Williamson,M.P. (1994) The structure and function of proline-rich regions in proteins. *Biochem. J.*, **297(Pt 2)**, 249–260.
41. Kay,B.K., Williamson,M.P. and Sudol,M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.*, **14**, 231–241.
42. Parikh,S.S., Walcher,G., Jones,G.D., Slupphaug,G., Krokan,H.E., Blackburn,G.M. and Tainer,J.A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA*, **97**, 5083–5088.
43. Zharkov,D.O., Mechetin,G.V. and Nevinsky,G.A. (2010) Uracil-DNA glycosylase: Structural, thermodynamic and kinetic aspects of lesion search and recognition. *Mutat. Res.*, **685**, 11–20.
44. Serrano-Heras,G., Bravo,A. and Salas,M. (2008) Phage phi29 protein p56 prevents viral DNA replication impairment caused by uracil excision activity of uracil-DNA glycosylase. *Proc. Natl Acad. Sci. USA*, **105**, 19044–19049.
45. Savva,R., McAuley-Hecht,K., Brown,T. and Pearl,L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.
46. Putnam,C.D., Shroyer,M.J., Lundquist,A.J., Mol,C.D., Arvai,A.S., Mosbaugh,D.W. and Tainer,J.A. (1999) Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with *Escherichia coli* uracil-DNA glycosylase. *J. Mol. Biol.*, **287**, 331–346.
47. Handa,P., Roy,S. and Varshney,U. (2001) The role of leucine 191 of *Escherichia coli* uracil DNA glycosylase in the formation of a highly stable complex with the substrate mimic, Ugi, and in uracil excision from the synthetic substrates. *J. Biol. Chem.*, **276**, 17324–17331.
48. Lundquist,A.J., Beger,R.D., Bennett,S.E., Bolton,P.H. and Mosbaugh,D.W. (1997) Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carboxylic amino acids in complex formation with *Escherichia coli* uracil-DNA glycosylase. *J. Biol. Chem.*, **272**, 21408–21419.
49. Tabor,S. (2001) Expression using the T7 RNA polymerase/promoter system. *Curr. Protoc. Mol. Biol.*, Chapter 16, Unit 16.2.