Published online 2 September 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 22* ***9779–9788***
doi:10.1093/nar/gkr667

Novel dimeric structure of phage ϕ29-encoded protein p56: insights into uracil-DNA glycosylase inhibition

Juan Luis Asensio[1,*], Laura Pérez-Lago[2], José M. Lázaro[2], Carlos González[3], Gemma Serrano-Heras[4] and Margarita Salas[2,*]

[1]Departamento de Química Orgánica Biológica, Instituto de Química Orgánica General, CSIC, 28006 Madrid, [2]Instituto de Biología Molecular 'Eladio Viñuela' (CSIC), Centro de Biología Molecular 'Severo Ochoa' (CSIC-UAM), Universidad Autónoma, Cantoblanco, 28049 Madrid, [3]Instituto de Química-Física 'Rocasolano', CSIC, 28006 Madrid and [4]Experimental Research Unit, General University Hospital of Albacete, C/ Hermanos Falcó, 37, 02006 Albacete, Spain

Received July 5, 2011; Revised and Accepted July 28, 2011

ABSTRACT

Protein p56 encoded by the *Bacillus subtilis* phage ϕ29 inhibits the host uracil-DNA glycosylase (UDG) activity. To get insights into the structural basis for this inhibition, the NMR solution structure of p56 has been determined. The inhibitor defines a novel dimeric fold, stabilized by a combination of polar and extensive hydrophobic interactions. Each polypeptide chain contains three stretches of anti-parallel β-sheets and a helical region linked by three short loops. In addition, microcalorimetry titration experiments showed that it forms a tight 2:1 complex with UDG, strongly suggesting that the dimer represents the functional form of the inhibitor. This was further confirmed by the functional analysis of p56 mutants unable to assemble into dimers. We have also shown that the highly anionic region of the inhibitor plays a significant role in the inhibition of UDG. Thus, based on these findings and taking into account previous results that revealed similarities between the association mode of p56 and the phage PBS-1/PBS-2-encoded inhibitor Ugi with UDG, we propose that protein p56 might inhibit the enzyme by mimicking its DNA substrate.

INTRODUCTION

Damage to DNA arises continually throughout the cell cycle and must be recognized and repaired prior to the next round of replication to maintain the genomic integrity of the cell. Uracil is one of the most common lesions in DNA (1,2). If left unrepaired, this change may impair protein–DNA interactions (3) or can induce G:C to A:T transition mutations during subsequent rounds of DNA replication (4). Uracil-DNA glycosylase (UDG) specifically recognizes uracil in DNA and initiates the base-excision repair (BER) mechanism by hydrolysing the N–C1′ glycosidic bond linking the uracil to the deoxyribose. This creates an abasic site that is removed by a 5′-acting apurinic/apyrimidinic (AP) endonuclease, leaving a gap that is filled by DNA polymerase and closed by DNA ligase. Four distinct families of UDGs have been identified in most prokaryotic and eukaryotic cells, showing very limited sequence similarity to each other, although family-1, the most ubiquitous, and family-2 proteins have been reported to possess the same structural fold (5,6). In addition, some DNA viruses, such as herpesviruses and poxviruses, also encode a UDG activity, whereas the human immunodeficiency virus type 1 packages cellular UDG (UNG2 enzyme) into virus particles. In these cases, the UDG activity appears to have an important role in virus replication (7,8).

Bacteria in their natural environment are faced with predation by both macro- and microorganisms. Some of

*To whom correspondence should be addressed. Tel: +34911964675; Fax: +34913978490; Email: msalas@cbm.uam.es
Correspondence may also be addressed to Juan Luis Asensio. Tel: +34915618806; Fax: +34915644853; Email: iqoa110@iqog.csic.es
Present address:
Laura Pérez-Lago, Servicio de Microbiología Clínica y Enfermedades Infecciosas, Hospital Gregorio Marañón, 28007 Madrid, Spain.

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

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

the more important of the predators are the bacteriophages. They have evolved different means of adapting to their host cells. For instance, it is known that several phages synthesize unique proteins that block critical cellular processes. These include proteins able to inhibit uracil-DNA repair of the host bacteria. The phage PBS-1/PBS-2 inhibitor Ugi represents the first example of such proteins (9–17). Ugi inactivates UDG by forming an irreversible 1:1 complex with the enzyme. The structural bases for this inhibitory activity have been analysed employing both NMR and X-ray methods. According to these data, complex formation is accompanied by a remarkable conformational change in the inhibitor, leading to a significant shape and electrostatic complementarity between interacting surfaces. Interestingly, the particular contacts observed at the UDG–Ugi interface suggest that Ugi achieves its tight binding by acting as a DNA mimic (12,16,17).

Recently, we reported the identification of a novel low molecular weight (56 amino acids) acidic inhibitor of the *Bacillus subtilis* UDG (18), called p56, which is encoded by the *B. subtilis* phage ϕ29. We have proposed that p56 constitutes a defence mechanism to prevent the deleterious effect caused by UDG due to elimination of uracil residues that may be present in the ϕ29 genome (19). In recent years, UDGs have emerged as attractive therapeutic targets due to their role in a wide range of biological processes including the generation of antibody diversity, DNA replication in a number of viruses and the formation of DNA strand breaks during anti-cancer drug therapy. Consequently, the identification and characterization of new molecules able to inhibit the activity of particular UDGs has a great interest. Indeed, some synthetic inhibitors of UDG have been designed to inhibit the human UNG enzyme (20). Herein, we address the structural bases for UDG inhibition by protein p56, combining the NMR structural study of p56 with the functional analysis of specific single and multiple mutants of the inhibitor.

MATERIALS AND METHODS

DNA substrates

To test the UDG activity, 34-mer oligonucleotides containing a single uracil (U) residue at position 16 (ssDNA-U16) (obtained from Isogen Bioscience BV) were used as ssDNA substrate. They were 5′-labelled with [γ-^{32}P] ATP (3000 Ci/mmol) (Perkin Elmer Life Science) and phage T4 polynucleotide kinase (New England Biolabs), and purified electrophoretically on 8 M urea/20% polyacrylamide gels. To generate dsDNA substrates, the 5′-^{32}P-labelled oligonucleotides were annealed to complementary non-labelled oligonucleotides (34-mer), containing either a guanine or adenine residue opposite to uracil in a buffer containing 20 mM Tris–HCl, pH 8.0 and 60 mM NaCl, heating at 70°C for 10 min and then slowly cooling to room temperature.

Construction of the expression plasmids

Gene *ung* of *B. subtilis*, which encodes UDG, was cloned in the *Escherichia coli* expression vector pGEX-2T (GST Gene Fusion System, GE Healthcare) and the resulting plasmid pGEX-2T-UDG wt was expressed in *E. coli* BL21 cells as described (21). Plasmid pT7-3-p56-ϕ29, containing the p56-encoding gene of phage ϕ29 was constructed as described (21).

Site-directed mutagenesis of protein p56

The p56 mutants were obtained by using the QuickChange site-directed mutagenesis kit obtained from Stratagene. Vector pT7-3-p56 containing phage ϕ29-p56 gene, was used as template for the mutagenesis reaction. After PCR using *Pfu* turbo DNA polymerase and treatment with DpnI endonuclease, amplified DNA was transformed into *E. coli* BL21(DE3) competent cells. The presence of the desired mutations and the absence of non-desired ones were confirmed by sequencing the entire gene.

Overexpression and purification of UDG and p56 proteins

Bacillus subtilis UDG was expressed and purified as described (21). Protein p56, wild-type and mutants, were overproduced in *E. coli* BL21(DE3) cells harbouring plasmids with the corresponding p56 gene and were purified as described (21). Purified proteins were quantified by gel densitometry and stored at −70°C.

Determination of UDG activity

UDG activity and the effect of protein p56 on the UDG activity were determined as described (21).

Sedimentation through glycerol gradients

Samples containing a mixture of lysozyme (15 μg), aprotinine (15 μg) and p56 (12 μg) proteins were loaded on top of a continuous 15–30% (v/v) glycerol gradient (4 ml) in the presence of 50 mM Tris–HCl, pH 7.5, 0.2 M NaCl, 1 mM EDTA and 7 mM β-mercaptoethanol, and centrifuged at 4°C for 49 h at 59 000 rpm in a Beckman TST 60.4 rotor. Gradients were fractionated and subjected to SDS–10–20% PAGE. Gels were stained after electrophoresis with Coomassie blue.

Isothermal titration microcalorimetry experiments

Binding studies were performed at 25°C in 10 mM sodium phosphate, pH 6.5 and 100 mM NaCl using a VP-ITC titration calorimeter (MicroCal, LLC) with a reaction cell volume of 1.467 ml. Both, enzyme and inhibitor solutions were dialysed against the same buffer prior to Isotermal Titration Microcalorimetry (ITC) experiments to ensure chemical equilibration. A solution 13 μM of UDG in the reaction cell was titrated with a 300 μM solution of p56 (monomer concentration) contained in a 300-μl syringe. At least 30 consecutive injections of 8 μl were applied at 5-min intervals, while the UDG solution was stirred at a constant speed of 300 rpm. Dilution heats of p56 into UDG solution (which agreed with those obtained by injections of the protein into the same volume of buffer) were subtracted from measured heats of binding. Titration curves were analysed with Origin, provided with the instrument by MicroCal LLC, using a one-site binding model to fit the curves. Thermodynamic

parameters were derived from two independent experiments and averaged.

NMR experiments

^{1}H-NMR spectra were recorded in 90:10 $^{1}H_2O$:$^{2}H_2O$ on Bruker Avance 800, Bruker AMX-600 and Varian Unity 500 MHz spectrometers. Protein concentrations for the structural analysis of the protein dimer were in the 400–800 μM range in, 10 mM sodium phosphate, pH 5.0 and 100 mM NaCl. Protein assignments were obtained using a set of 2D- and 3D homo- and heteronuclear NMR experiments performed at 35°C on the unlabelled, ^{15}N-labelled and ^{13}C/^{15}N double-labelled molecules. Thus, HNCACB and CBCA(CO)NH spectra were employed for backbone assignment. NOE distance restraints were obtained from ^{15}N- or ^{13}C-edited 3D-NOESY spectra. In addition, 2D-NOESY, ^{13}C-selected-^{12}C-filtered experiments were performed on a heterolabelled dimer to analyse the interprotein contacts. This protein sample was generated by mixing equimolar amounts of unlabelled and ^{13}C/^{15}N double-labelled protein to a global protein concentration of 100 μM in 10 mM phosphate buffer, pH 5.0 and 100 mM NaCl. The sample was incubated at 40°C for 24 h and then concentrated for NMR analysis.

Upper limits for proton–proton distances were obtained from NOESY cross-peak intensities at three mixing times, 50, 75 and 150 ms. Cross-peaks were classified as strong, medium, and weak corresponding to upper limits of 2.6, 3.5 and 5.0 Å. The lower limit for proton–proton distances was set as the sum of the van der Waals radii of the protons. Structure calculations were performed using the program DYANA (22). A set of 1380 constraints (120 interprotein) were used in the final round of calculations. The 30 best DYANA structures in terms of target function were subjected to a simulated annealing protocol with the AMBER 10.0 package (23,24).

RESULTS AND DISCUSSION

Stoichiometry of the p56–UDG complex determined by microcalorimetry

By analytical ultracentrifugation, it has been shown that p56 is dimeric under physiological conditions (25). To determine the stoichiometry of the p56–UDG complex, the binding of the inhibitor to the enzyme was analysed by Isothermal Titration Microcalorimetry (ITC). A typical titration profile together with the *K*b, ΔG and ΔH parameters derived from titration experiments, assuming a single set of binding sites in the enzyme, is shown in Figure 1. According to these data, one p56 dimer binds to a single UDG molecule to form a 2:1 p56–UDG complex. From a structural perspective, the observed stoichiometry is intriguing and strongly suggests that a significant fraction of the surface of the inhibitor dimer must be occluded by the enzyme upon binding so that the association of a second UDG molecule to form a symmetric 2:2 complex is impeded. It should be noted that this implies that p56 dimerization is probably essential for complex formation (this was confirmed by the functional analysis of specific p56 mutants. See below).

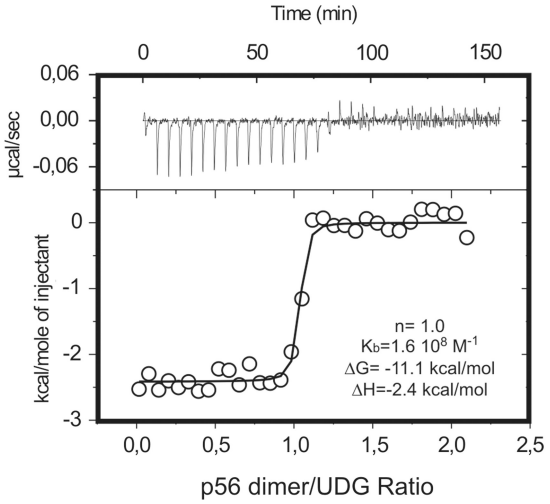

Figure 1. Stoichiometry of the p56–UDG interaction as deduced from isothermal titration calorimetry (ITC). UDG solutions were titrated with p56 inhibitor at 25°C in 10 mM sodium phosphate, pH 6.5 and 100 mM NaCl. The experimental curves were fitted assuming a single set of sites on the enzyme. The obtanied stoichiometry factor (*n*), binding constant (*K*b), binding free energy (ΔG) and enthalpy (ΔH) values are shown.

Regarding the stability of the complex, the combination of very high affinity and low enthalpy prevented an accurate measurement of *K*b, which seems to be in the $10^8\,M^{-1}$ range. Thus, the unusually low enthalpy value detected for such tight interaction might point to a dominant role of electrostatic forces in the stabilization of the enzyme–inhibitor complex. These data are consistent with our previous findings showing that the interaction between UDG and p56 is stable at high urea (6 M) concentrations (25).

Architecture of the p56 dimer

To gain insights into the mechanism underlying the UDG inhibition by protein p56, we analysed the solution structure of the inhibitor by NMR. Backbone and side-chain assignments (Supplementary Figure S1 and Table S1) of p56 were obtained using standard 2D and 3D NMR techniques on ^{15}N- and ^{13}C/^{15}N-labelled samples. In order to distinguish intra- from intermolecular NOEs, half-filtered NOESY experiments were carried out on a heterolabelled dimer formed by mixing equivalent amounts of unlabelled and ^{13}C/^{15}N-labelled protein. A summary of the experimental constraints employed and the characterization of the final NMR ensemble is shown in Table 1 (see also Supplementary Table S2).

The solution structure of the p56 dimer is shown in Figure 2, together with the electrostatic potential at the protein surface. In addition, details of the dimerization inter-phase are presented in Figure 3. According to the NMR data, each polypeptide chain contains three stretches of anti-parallel β-sheets (β_1: residues 10–17; β_2: residues 22–29; β_3: residues 48–54) and a helical region (α_1: residues 33–43) linked by short loops (L_1: residues 18–21; L_2: residues 30–32 and L_3: residues 44–47). The

Table 1. Characterization of the 20 NMR structures of the p56 dimer retained for structural analysis

NMR distance constraints	
Total distance constraints	1380
Intraresidue	216
Sequential	268
Medium and long range	896
Intra-monomer	1260
Inter-monomer	120
RMSD (residues 8–55, 20 structures)	
AMBER backbone RMSD (dimer)	1.17 ± 0.24
AMBER heavy atom RMSD (dimer)	1.88 ± 0.21
AMBER backbone RMSD (monomer)	1.02 ± 0.28
AMBER heavy atom RMSD (monomer)	1.77 ± 0.26
RMSD (residues 8–17, 23–43, 46–55, 20 structures)	
AMBER backbone RMSD (dimer)	0.82 ± 0.13
AMBER heavy atom RMSD (dimer)	1.60 ± 0.14
AMBER backbone RMSD (monomer)	0.76 ± 0.15
AMBER heavy atom RMSD (monomer)	1.54 ± 0.20
Distance constraints violations (>0.2 Å)	
Sum of violations	1.41 (1.11–1.69)
Maximum violation	0.17 (0.14–0.21)
AMBER energies (kcal/mol)	
Restraint AMBER energy	7.5 (5.4–9.2)
Total AMBER energy	−1476 (−1255 to −1764)

RMSD values and violations are given in angstroms.

connectivity of secondary structural elements is β_1–β_2–α_1–β_3. It can be observed that helix α1 sits on strand β1 with a parallel orientation, forming a prominent and extensive hydrophobic core. Numerous NOE contacts connecting these protein regions were observed in both 2D and 3D experiments.

While the overall structure appears well defined (Figure 2A and Table 1), the NMR ensemble exhibits a significant dispersion at four distinct regions of the polypeptide: the N-terminal segment (residues 1–7), the C-terminal end of helix α_1 (residues 41–43) and loops L_1/L_3 (comprising residues 18–21 and 44–47, respectively). The lack of structural definition at the helical segment is probably related with the presence of a glycine (a residue with a very low helical propensity) at position 41. Indeed, the medium-range NOEs at this protein region are consistent with a significantly reduced helical content. In a similar way, diverse experimental evidences (as the extreme line broadening observed for specific proton signals in residues 44 and 47) indicate that, at least for the N-terminus and loop L_3, the lack of structural definition is caused by substantial local motions of the protein backbone.

Resolution of the crystal structure of phage PBS-2 inhibitor Ugi revealed that this protein is essentially an anti-parallel β-sheet made up of five contiguous strands with a helix in the polypeptide stretch that connects strands 1 and 2 (16,17). It should be noted that, despite the existing similarities in secondary structure between p56 and Ugi, both protein inhibitors exhibit a fundamentally different fold (Figure 2C). Indeed, a DALI (26) search reveals that p56 defines a novel dimeric structure, not having high structural homology with any of the proteins present in the Brookhaven Protein Data Bank. It can be observed that the two monomers, related by a 2-fold symmetry axis, are held together by a combination of hydrogen bonding, and extensive hydrophobic and electrostatic interactions. Two specific regions of p56 seem to play a key role in the stabilization of the structure: strand β_3 and the helical segment α1 (Figures 2, 3, Supplementary Figure S2 and Table S2). In the following description of the main intermolecular contacts, letters **a** and **b** will be employed as subscript to specify the polypeptide chain only when essential to avoid ambiguities. Thus, strands β_3 of both monomers (β_{3a} and β_{3b}) extend across the dimer interface in a 2-fold symmetric fashion to create a continuous six-stranded anti-parallel β-sheet. This particular arrangement implies that the two helices, α_{1a} and α_{1b}, lye anti-parallel to one another on the same side of the β-sheet. Residues on strands β_{3a}/β_{3b} with side chains pointing to the helical segments α_{1a}/α_{1b} are mainly aliphatic and participate in numerous intermolecular interactions. Most relevant contacts at this region involve pairs $L33_a$/$I47_b$, $L50_a$/$L50_b$ and $I47_a$/$L33_b$ (Figure 3B). In addition, helices α_{1a}/α_{1b} contribute to the stabilization of this hydrophobic core through residues L33, F36 and Y40, which are also involved in numerous intra- and intermolecular packing interactions. Intriguingly, two acidic amino acids, $E37_a$ and $E37_b$, are located at the centre of the helical segments, pointing to the dimerization interface. It seems unlikely that these residues have a favourable influence on the protein stability, especially considering the proximity between both charged side-chains within the dimer structure and the hydrophobic nature of most of the surrounding positions. Finally, the water-exposed face of strand β_3 is characterized by a significant fraction of charged residues. More specifically, this region seems to contribute to the stabilization of the p56 dimer by means of two, symmetry related, bifurcated salt-bridges involving triads $K49_a$/$D53_b$/$R51_b$ and $K49_b$/$D53_a$/$R51_a$.

Protein p56 dimerization is required for the inhibition of UDG

The NMR structure of p56 provides a detailed description of the particular interactions involved in the stabilization of the protein dimer. This information was used to analyse the role that p56 dimerization plays on UDG inhibition. More specifically, we investigated whether dimer formation is required to inhibit the UDG activity. To this end, several residues located at the protein–protein interface of the inhibitor were substituted by Ala, and the influence of the mutations on both the dimer stability and its inhibitory activity were evaluated. The mutated positions included hydrophobic and polar side-chains located at both sides of the dimerization interface (formed by helical segments α_{1a}/α_{1b} and strands β_{3a}/β_{3b}). In particular, single mutants Y40A and D53A, and the double mutant K49A/R51A were constructed and purified as described in 'Materials and Methods' section. According to the NMR structure of p56, Y40 is partially buried in an extensive hydrophobic core defined by the packing of both p56 monomers. In contrast, D53, K49 and R51 are located on the water exposed surface of the six-stranded

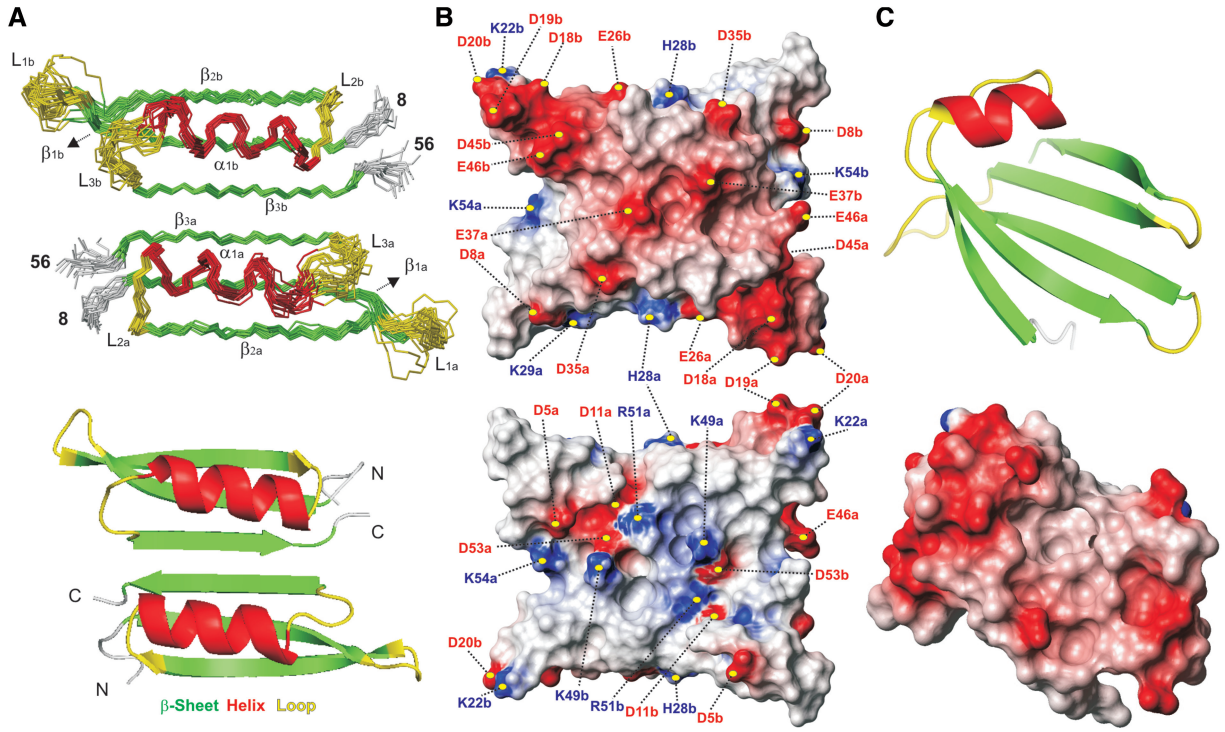

Figure 2. (**A**) Backbone trace for the 20 lowest energy structures in the final NMR ensemble of the p56 dimer (Top) together with a ribbon representation of a single conformer (Bottom). Different monomeric units are labelled with subscripts **a** and **b**. Secondary structure elements are indicated by colour coding: α-helices (α_1, red), β-strands (β_1–β_3, green), and loops (L_1–L_3, yellow). The disordered N-terminal tails (residues 1–7) are omitted for clarity. (**B**) Molecular surface representations of p56 showing the electrostatic potential (blue is positive and red is negative). The orientation of the p56 dimer in the top view is identical to that of the NMR ensemble shown in (**A**). Bottom view is related with the former by a 180° rotation around the *x*-axis. The position of the charged residues contributing to the inhibitor electrostatic potential is indicated. (**C**) Top: ribbon representation of the UDG-bound form of the inhibitor Ugi from phage PBS-1/PBS-2 (16) (pdb code 1EUI). Bottom: molecular surface representation showing the electrostatic potential at the UDG-binding surface of Ugi inhibitor.

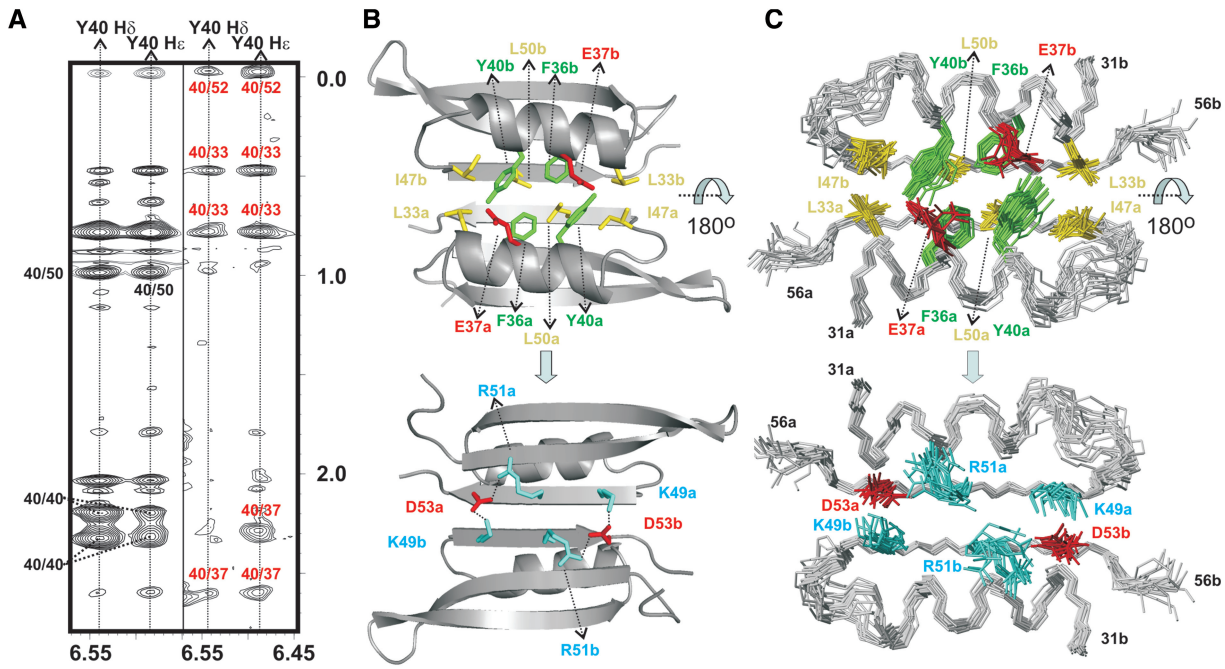

Figure 3. (**A**) Double-filtered (left) and half-filtered (right) NOESY experiments acquired for p56 at 35°C (buffer conditions: 100 mM NaCl and 10 mM sodium phosphate, pH 5.0). Inter-protein NOE contacts involving the aromatic ring of Y40 and residues E37, L33 and L52 are clearly observable in the half-filtered spectrum (labelled in red). (**B**) Structural details of p56 dimerization interface. Protein side-chains involved in relevant protein–protein contacts are highlighted (aromatic, in green; aliphatic, in yellow; positively charged, in cyan; negatively charged, in red). The protein backbone is shown in grey. Different monomeric units are labelled with subscripts **a** and **b**. Top and bottom views are related by a 180° rotation around the *x*-axis. (**C**) Ensemble of 20 NMR structures showing the inhibitor side-chains involved in dimerization.

β-sheets formed by the p56 dimer, and participate presumably in stabilizing bifurcated salt-bridges (Figure 3). It is noteworthy that this particular choice (mutated residues at buried sites or exposed on the positively charged surface of the protein) strongly reduces the probability that direct p56–UDG contacts are removed by the substitutions leading to a drop in inhibitory activity.

The oligomerization state of p56 mutants was analysed by glycerol gradient sedimentation. As shown in Figure 4A, p56 wild-type sedimented in glycerol gradients near the position of lysozyme, which is in agreement with its dimeric state. In contrast, peaks corresponding to Y40A, D53A or K49A/R51A mutants were displaced towards the top of the gradient, indicating that such mutants failed to assemble into dimers. This result is fully consistent with the structural information and shows that the selected positions play a key role in the stabilization of the p56 dimer. Then, the ability of the dimerization mutants to inhibit the UDG activity was determined. To this end, the 34-mer oligonucleotide containing a single uracil was incubated with 5 pg of UDG in the presence of different amounts of p56 mutants (from 0.2 to 5.4 ng). Interestingly, these experiments revealed a significant drop of the inhibitory activity for all the mutants (Figure 4B and C). More specifically, the capacity of mutants Y40A and D53A to block the UDG activity was decreased 18- and 14-fold, respectively. The double mutant was less affected, though it only retained 20% of its inhibition activity. The higher activity exhibited by the K49A/R51A mutant with respect to the D53A mutant might reflect the increased anionic character of the former. Overall, the results obtained strongly suggest that the two monomers of p56 participate in contacts with UDG. In addition, it should be considered that monomeric p56 would probably exhibit both an enhanced internal mobility and a reduced stability, which could have a significantly unfavourable influence on its binding properties. Indeed, p56 solution structure strongly suggests that folding and dimerization are strongly coupled. Altogether our data indicate that the formation of p56 dimer is a necessary step for inhibition of UDG.

Effects of mutations in acidic residues on the inhibition activity of p56

It has been shown that the tight binding of the protein inhibitor Ugi to human UDG arises from its ability to functionally mimic substrate DNA (12). In agreement with this, Ugi regions involved in UDG recognition exhibit a strongly negative electrostatic potential (Figure 2C). Recent studies have shown that Ugi, which interacts with the DNA-binding groove of the enzyme, replaces the highly acidic protein p56 previously bound to *E. coli* UDG, implying that both inhibitors present at least partially overlapping binding sites on the enzyme (25). Moreover, it has also been shown that p56 competes with DNA for binding to UDG and that the formation of UDG–p56 complex blocks the interaction between the enzyme and DNA (25). Altogether, these results suggest that, the p56 surface involved in UDG recognition should be strongly anionic. Interestingly, acidic amino acids are present at equivalent positions in most p56-like proteins (21), which might suggest a possible role for them in UDG inhibition. Furthermore, as shown here, the NMR solution structure of p56 revealed that most aspartates and glutamates fall in the same face of the dimer, specifically, at the surface defined by helices $\alpha1_a/\alpha1_b$ and loop regions L_{1a}/L_{1b} and L_{3a}/L_{3b} (Figure 2B). Indeed, triads $D18_a/D19_a/D20_a$ and $D18_b/D19_b/D20_b$, present in the protruding loops L_{1a}/L_{1b}, together with pairs $D45_a/E46_a$ and $D45_b/E46_b$, and residues $E37_a/E37_b$, confer to these regions a strongly negative potential. On the contrary, the opposite side of p56, defined by the water exposed surface of the six-stranded β-sheet exhibits a moderate positive potential, which results from the presence of up to eight basic side-chains of $K22_a/K22_b$, $K49_a/K49_b$, $R51_a/R51_b$ and $K54_a/R54_b$. Hence, according to this analysis, helical regions α_{1a}/α_{1b} and loops L_{1a}/L_{1b} and L_{3a}/L_{3b} constitute the most likely UDG-binding surface of p56.

To test this hypothesis, we analysed the role of specific p56 acidic residues in the UDG inhibition. In particular, we focused our attention on glutamate/aspartate side-chains clustered on particular regions of the inhibitor and also on those conserved in most p56-like proteins from other ϕ29-related phages (21). Thus, we constructed two single mutants at positions 26 and 37, in which conserved acidic residues were replaced by alanine. In addition, acidic residues clustered on loop L_1 were neutralized by the triple mutation D18A/D19A/D20A. These positions, together with E37, contribute considerably to the negative electrostatic potential exhibited by the proposed p56 UDG-binding surface. Finally, we generated the triple mutant D5A/D8A/D11A, with substitutions in the N-terminal tail of the protein. It should be noted that, with the exception of E37, all the residues considered are water exposed and far from the dimerization interface. Consequently, it is unlikely that the selected mutations have a significant influence on the protein stability or structure. As shown in Figure 5A, a strong inhibition of uracil excision by UDG was observed in the presence of 0.8 ng of wild-type p56. For the mutant D18A/D19A/D20A, a 16-fold increase in the concentration of the inhibitor was required to detect the same effect (Figure 5C). In contrast, the triple mutant D5A/D8A/D11A exhibited a modest 2-fold decrease in the inhibition activity. Regarding the substitutions at conserved positions 26 and 37, the mutants were much less efficient than the wild-type p56 at preventing the cleavage of uracil-containing substrate by UDG, especially mutant E37A (Figure 5A and B). More specifically, E26A and E37A lost 87% and 96%, respectively, of inhibition activity with respect to the wild-type protein (Figure 5C).

Therefore, our results confirm the key role played by acidic residues present in loop L_1 and at positions 26 and 37 in UDG inhibition, likely due to the fact that these amino acids are involved in binding to the UDG. This is consistent with mutational analysis studies carried out with protein Ugi, which showed the role of electrostatic interactions between the specific acidic residues of the β_1 edge of the inhibitor with the active site residues of *E. coli* UDG in the formation UDG–Ugi complex (15,17).

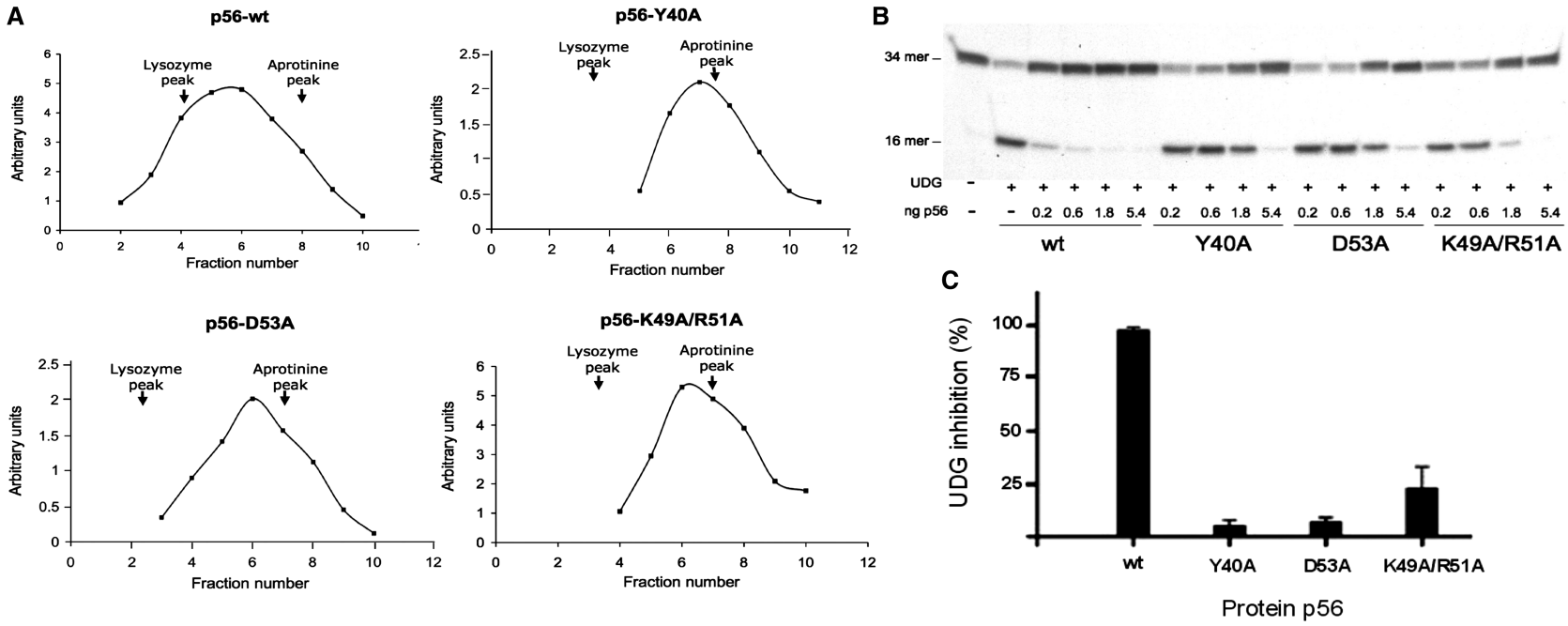

Figure 4. Influence of selected residues on the stability of the p56 dimer and its inhibition activity. (**A**) The oligomerization state of p56 mutants was analysed by glycerol gradient sedimentation. Wild-type p56 and mutants, Y40A, D53A and K49A/R51A (12 μg of each), were loaded on a 15–30% glycerol gradient and subjected to centrifugation for 49 h at 59 000 rpm in a Beckman SW.65 rotor. As markers, 15 μg of lysozyme (14.7 kDa) and 15 μg of aprotinine (6.5 kDa) were loaded in the same gradient. Sedimentation was from right to left. After fractionation, aliquots from each fraction were analysed by SDS–PAGE. Densitometry scanning of the gels stained with Coomassie Blue was used to determine the amount of the different proteins in each fraction (arbitrary units). The arrows indicate fractions at which the maximal amount of each marker was detected. (**B**) Inhibition activity of p56 dimerization mutants. Increasing amounts of the p56 mutants (from 0.2 to 5.4 ng) were incubated with 5 pg of UDG. Then, the DNA substrate was added and incubated for 20 min at 37°C and the product was analysed in 8 M urea–20% polyacrylamide gels. (**C**) Bands corresponding to DNA substrate and DNA product were quantified by densitometry and the percentage of inhibition was calculated taking the wild-type p56 values as 100% inhibition. Data are depicted in a bar chart and average values of three independent experiments with standard deviations are represented.

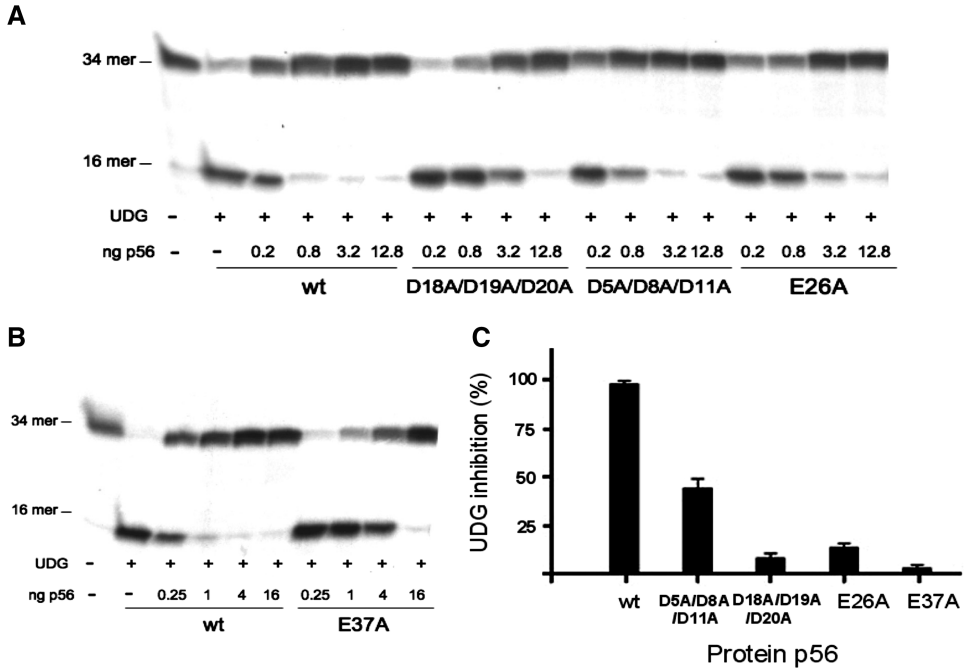

Figure 5. Influence of acidic residues on the p56 inhibition activity. The inhibition ability of two triple mutants (D5A/D8A/D11A and D18A/D19A/D20A) and two single mutants, E26A(**A**) and E37A(**B**) of p56 was assayed. Different doses of the mutant proteins (from 0.2 to 18.8 ng for the triple mutants and the E26A mutant and from 0.25 to 16 ng for the E37A mutant) were incubated with 5 pg of UDG. Then, the DNA substrate was added and incubated for 20 min at 37°C and the product was analysed in 8 M urea-20% polyacrylamide gels. (**C**) After quantification by densitometry, the percentage of inhibition for the mutants was calculated with respect to the wild-type protein. Data are depicted in a bar chart and average values of three independent experiments with standard deviations are represented.

However, although both p56 and Ugi contain negatively charged residues required for the interaction with UDG, these inhibitors exhibit a clearly different protein structure and distinct binding stoichiometries. On the other hand, the findings suggest that p56 N-terminal region plays a modest role in the inhibition, maybe contributing to the stabilization of the p56–UDG complex by long-range non-specific electrostatic interactions. In fact, negative electrostatic field present in proteins that mimic DNA, such as transcription factor $dTAF_{II}230$ from *Drosophila* and bacteriophage T7 Ocr protein, has been reported to facilitate productive collisions between such proteins and their targets (27–29).

CONCLUSION

To gain insights into the UDG inhibition mode by protein p56, the NMR solution structure of the inhibitor has been determined. According to these data, p56 defines a novel dimeric structure characterized by a 2-fold symmetry axis. The protein dimer seems to be stabilized by a combination of polar and extensive hydrophobic interactions involving mainly residues located on helices α_{1a}/α_{1b} and β-strands β_{3a}/β_{3b}. The relevance of these positions in p56 dimerization has been shown by site-directed mutagenesis. Furthermore, the functional analysis of specific mutants of the protein, unable to assemble into dimers, confirmed that this process is essential for UDG inhibition, as deduced from the microcalorimetry assays.

On the other hand, p56 solution structure also revealed the presence of a large solvent exposed surface characterized by a strongly negative electrostatic potential. This comprises the helical segments, together with loops L_1 and L_3. The involvement of these protein regions in UDG recognition was shown by the analysis of several single and multiple mutants of p56, in which acidic residues clustered at these sites or conserved in p56-like proteins from other ϕ29-related phages, were replaced by alanines.

In summary, our results indicate that p56 and Ugi, display a significantly different fold. Indeed, p56 defines a novel dimeric structure that binds to UDG in a 2:1 stoichiometry. It has been shown that Ugi binding to UDG is accompanied by a remarkable conformational change of the inhibitor, involving a substantial reorientation of two helical segments. Interestingly, the NMR analysis of Ugi in the free state has revealed that both helices project away from the β-sheet core, and therefore are located on potentially flexible arms of the polypeptide. This specific property would greatly reduce the energy cost of the observed adjustment. In contrast with Ugi, p56 exhibits a more compact structure. Indeed, all the secondary structure elements participate in tight packing interactions, which seem essential to preserve the structural integrity of the inhibitor. Consequently, it is highly unlikely that major structural rearrangements (as those described for Ugi) take place upon binding to UDG. Nevertheless, minor adjustments of the protein flexible regions

(in particular loops L1 and L3 and the N-terminal tail) would be certainly feasible, and might contribute to increase the charge and shape complementarity between enzyme and inhibitor.

Despite the structural differences existing between p56 and Ugi, several findings suggest that both proteins present similar inhibition modes. Previous studies showed that Ugi is able to replace protein p56 previously bound to *E. coli* UDG (25), and that UDG position 191 is required for the inhibition mediated by both inhibitors (21,30). Furthermore, the present article confirms the role of acidic residues of p56 in the UDG inhibition, as previously reported for the formation of UDG–Ugi complex (15). Thus, these observations suggest that protein p56 might function as a structural mimic of the DNA substrate recognized by UDG.

COORDINATES

Coordinates for the 20 final NMR structures of p56 dimer have been deposited in the Protein Data Bank with accession code 2LE2.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Spanish Ministry of Science and Innovation (CTQ2010-19073 to J.L.A., CTQ2010-21567-CO2-02 to CG and BFU2008-00215 and CONSOLIDER-INGENIO CSD2007-00015 to M.S.); Autonomous Community of Madrid (S2009MAT-1507 to M.S.); Fundación Mutua Madrileña (to M.S.); Fundación Ramón Areces to the Centro de Biología Molecular 'Severo Ochoa' Institutional grant. Funding for open access charge: Spanish Ministry of Science and Innovation (BFU2008-00215).

Conflict of interest statement. None declared.

REFERENCES

1. Lindahl,T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
2. Kavli,B., Otterlei,M., Slupphaug,G. and Krokan,H.E. (2007) Uracil in DNA-general mutagen, but normal intermediate in acquired immunity. *DNA Repair*, **6**, 505–516.
3. Mosbaugh,D.W. and Bennett,S.E. (1994) Uracil-excision DNA repair. *Prog. Nucleic Acids Res. Mol. Biol.*, **48**, 315–370.
4. Krokan,H.E., Drablos,F. and Slupphaug,G. (2002) Uracil in DNA–occurrence, consequences and repair. *Oncogene*, **21**, 8935–8948.
5. Pearl,L.H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.*, **460**, 165–181.
6. Aravind,L. and Koonin,E.V. (2000) The alpha/beta fold. Uracil DNA glycosylases: a common origin with diverse fates. *Genome Biol.*, **1**, RESEARCH 0007-research 0007.8.
7. Olsen,L.C., Aasland,R., Wittwer,C.U., Krokan,H.E. and Helland,D.E. (1989) Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *EMBO J.*, **8**, 3121–3125.
8. Chen,R., Wang,H. and Mansky,L.M. (2002) Roles of uracil-DNA glycosylase and dUTPase in virus replication. *J. Gen. Virol.*, **83**, 2339–2345.
9. 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.
10. Bennett,S.E. and Mosbaugh,D.W. (1992) Characterization of the *Escherichia coli* uracil-DNA glycosylase inhibitor-protein complex. *J. Biol. Chem.*, **267**, 22512–22521.
11. Bennett,S.E., Schimerlik,M.I. and Mosbaugh,D.W. (1993) Kinetics of the uracil-DNA glycosylase inhibitor-protein association. Ung interaction with Ugi, nucleic acids, and uracil compounds. *J. Biol. Chem.*, **268**, 26879–26885.
12. 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.
13. Balasubramanian,S., Beger,R.D., Bennett,S.E., Mosbaugh,D.W. and Bolton,P.H. (1995) Secondary structure of uracil-DNA glycosylase inhibitor protein. *J. Biol. Chem.*, **270**, 296–303.
14. Beger,R.D., Balasubramanian,S., Bennett,S.E., Mosbaugh,D.W. and Bolton,P.H. (1995) Tertiary structure of uracil-DNA glycosylase inhibitor protein. *J. Biol. Chem.*, **270**, 16840–16847.
15. 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.
16. 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.
17. 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.
18. 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.
19. Serrano-Heras,G., Bravo,A. and Salas,M. (2008) Phage ø29 protein p56 prevents viral DNA replication impairment caused by uracil excision activity of uracil-DNA glycosylase. *Proc. Natl Acad. Sci. USA*, **105**, 19044–19049.
20. Jiang,Y.L., Krosky,D.J., Seiple,L. and Stivers,J.T. (2005) Uracil-directed ligand tethering: an efficient strategy for uracil DNA glycosylase (UNG) inhibitor development. *J. Am. Chem. Soc.*, **127**, 17412–17420.
21. 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 ø29 protein p56. *Mol. Microbiol.*, **80**, 1657–1666.
22. Guntert,P., Mumenthaler,C. and Wuthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.
23. Pearlman,D., Case,D.A., Caldwell,W.J., Ross,W.R., Cheatham,T.E., DeBolt,S., Ferguson,D., Seibel,G. and Kollman,P. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal-mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.*, **91**, 1–41.
24. Case,D., Darden,T.A., Cheatham,T.E., Simmerling,C.L., Wang,J., Duke,R.E., Luo,R., Crowley,M., Walker,R.C., Zhang,W. *et al.* (2008) AMBER 10, University of California, San Francisco.
25. Serrano-Heras,G., Ruiz-Masó,J.A., del Solar,G., Espinosa,M., Bravo,A. and Salas,M. (2007) Protein p56 from the *Bacillus subtilis* phage ø29 inhibits DNA-binding ability of uracil-DNA glycosylase. *Nucleic Acids Res.*, **35**, 5393–5401.

26. Holm,L. and Sander,C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res.*, **26**, 316–319.
27. Liu,D., Ishima,R., Tong,K.I., Bagby,S., Kokubo,T., Muhandiram,D.R., Kay,L.E., Nakatani,Y. and Ikura,M. (1998) Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell*, **94**, 573–583.
28. Atanasiu,C., Byron,O., McMiken,H., Sturrock,S.S. and Dryden,D.T. (2001) Characterisation of the structure of ocr, the gene 0.3 protein of bacteriophage T7. *Nucleic Acids Res.*, **29**, 3059–3068.
29. Stephanou,A.S., Roberts,G.A., Tock,M.R., Pritchard,E.H., Turkington,R., Nutley,M., Cooper,A. and Dryden,D.T. (2009) A mutational analysis of DNA mimicry by ocr, the gene 0.3 antirestriction protein of bacteriophage T7. *Biochem. Biophys. Res. Commun.*, **378**, 129–132.
30. 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.