Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic studies of a complex between the *Lactococcus lactis* Fpg DNA-repair enzyme and an abasic site containing DNA

Karine Pereira de Jésus,[a] Laurence Serre,[b] Nadège Hervouet,[a] Véronique Bouckson-Castaing,[a]† Charles Zelwer[a] and Bertrand Castaing[a]*

[a]Centre de Biophysique Moléculaire UPR4301, affiliated to the University of Orléans, CNRS, Rue Charles Sadron, 45071 Orléans CEDEX 02, France, and [b]Institut de Biologie Structurale, CNRS–CEA, 41 Avenue Jules Horowitz, 38027 Grenoble CEDEX 01, France

† Present address: Professor at Lycee Ste Croix-St Euverte, 45000 Orléans, France.

Correspondence e-mail: castaing@cnrs-orleans.fr

Received 9 November 2001
Accepted 21 January 2002

For protein–DNA complex crystallization, the choice of the DNA fragment is crucial. With the aim of crystallizing the 31 kDa Fpg DNA-repair enzyme bound to DNA, oligonucleotide duplexes varying in length, sequence, end type and nature of the specific DNA target site were used. Crystals of several protein–DNA combinations grew from solutions containing both polyethylene glycol and salt. This systematic crystallization screening followed by optimization of the crystallization conditions by microseeding led to crystals of Fpg bound to a 13 base-pair duplex DNA carrying the 1,3-propanediol abasic site analogue which are suitable for crystallographic analysis. Complete native data sets have been collected to 2.1 Å resolution.

1. Introduction

The bacterial formamidopyrimidine-DNA glycosylase (Fpg or MutM) is one of the enzymes that initiate the base-excision repair pathway. First identified in *Escherichia coli* as a DNA glycosylase (hydrolysis of the N1–C1′ glycosydic bond) able to remove imidazole ring-opened purines (Fapy residues; Chetsanga & Lindahl, 1979), Fpg was subsequently reported to bind and remove a wide range of oxidatively damaged bases such as 8-oxoguanine (8-oxoG), one of the major oxidized purine products (Tchou *et al.*, 1991; Czeczot *et al.*, 1991; Hatahet *et al.*, 1994). Fpg also displays an AP (apurinic/apyrimidic) lyase activity which excizes the resulting AP site. This second reaction occurs through successive 3′- and 5′-phosphate cleavages by a $\beta\delta$-elimination mechanism, leading to a one-nucleoside gap in the DNA (Bailly *et al.*, 1989; O'Connor & Laval, 1989). After 3′-dephosphorylation, DNA polymerase can fill in the resulting short single-stranded gap and, finally, DNA continuity is restored by the DNA ligase (Lloyd & Lin, 1993). Thereby, Fpg prevents cell death and spontaneous G↔T transversions induced by replication arrests at Fapy residues (O'Connor *et al.*, 1989) and by the miscoding 8-oxoG, respectively (Michaels *et al.*, 1991; Castaing *et al.*, 1993).

fpg genes are highly conserved through a wide range of aerobic bacteria (Boiteux *et al.*, 1987; Duwat *et al.*, 1995; Mikawa *et al.*, 1998) and eukaryotic plants (Murphy & Gao, 1998; Ohtsubo *et al.*, 1998). These genes encode a 31 kDa monomeric enzyme containing a C-terminal zinc-finger motif (-Cys-X_2-Cys-X_{16}-Cys-X_2-Cys-) required for DNA binding (O'Connor *et al.*, 1993; Castaing *et al.*, 1993) and a strictly conserved N-terminal sequence (Pro-Glu-Leu-Pro-Glu-Val-Glu-Thr-) containing the catalytic Pro1. Fpg is a DNA glycosylase/AP lyase distinguished from monofunctional DNA glycosylases by its ability to form a covalent Schiff-base intermediate between the C1′-aldehyde function of the AP site and the α-amino group of the N-terminal proline during the AP lyase process (Zharkov *et al.*, 1997).

Recently, the crystal structure of the *Thermus thermophilus* Fpg (TtFpg) was determined at 1.9 Å resolution (Sugahara *et al.*, 2000). TtFpg is composed of two domains with original folds connected by a flexible hinge. Most of the conserved amino-acid residues in the Fpg family are clustered in a large electropositive cleft formed by the interface between the two Fpg domains. The conserved catalytic N-terminal proline emerges at the bottom of the cleft.

To understand how the protein locates, binds and processes its DNA substrate, we initiated the crystallographic study of abortive complexes between the Fpg protein from *Lactococcus lactis* (LlFpg) and DNA fragments containing AP-site analogues such as 1,3-propanediol and tetrahydrofuran. These DNA fragments are not substrates for the Fpg AP lyase activity, but constitute high-affinity ligands (Castaing *et al.*, 1999).

In this paper, we describe the crystallization strategy which allowed us to obtain crystals of LlFpg–DNA abortive complex diffracting to 2.1 Å resolution.

2. Experimental and results

2.1. Protein and DNA preparations

The stability of macromolecules is critical in crystallographic studies. Unfortunately, at high

concentrations (up to 3 mg ml^{-1}) and 277 K, the purified Fpg protein is slowly cleaved into two inactive peptides. However, this intrinsic instability is more pronounced in the *E. coli* enzyme (EcFpg) than in its *L. lactis* homologue (LlFpg). Moreover, we have also observed that P1G-LlFpg (a mutant enzyme with a mutation of Pro1 to glycine in LlFpg) is more stable than the wild-type LlFpg and is still able to bind (B. Castaing, unpublished observations) and to partially process damaged DNA (Zharkov *et al.*, 1997; Sidorkina & Laval, 2000). Considering these observations, we decided to use P1G-LlFpg for screening the initial crystallization conditions of Fpg–DNA complexes.

To obtain the P1G-LlFpg mutant protein, amino-acid replacement was achieved by PCR using the plasmid pVE1064 as a template (Duwat *et al.*, 1995) and the oligonucleotide primers TCACCCGGGGAATTACCAGAAGTTGAACCG and CTGATCTAGATTATTTTTGCTGACAAACTGG. After *Sma*I and *Xba*I digestions, the resulting insert was ligated into the pMAL-c vector (New England Biolabs), previously linearized by *Xba*I and *Xmn*I digestions. After sequence checking, the recombinant plasmid (pMAL-LP1G) was used to transform TG1 cells. LB medium containing 100 µg ml^{-1} ampicillin was inoculated with an overnight pre-culture of TG1 transformant cells. After IPTG induction (0.5 m*M*), P1G-LlFpg was overproduced as a fusion protein between maltose-binding protein (MBP) and P1G-LlFpg (MBP-P1G-LlFpg). Bacteria were then lyzed by several freeze-thaw cycles in the presence of lyzosyme (0.5 mg ml^{-1}). After ultracentrifugation clarification of the bacterial lysate, MBP-P1G-LlFpg was isolated from the soluble fraction (Fig. 1, lane I) by a series of chromatographic steps consisting of QMA (Waters) (lane II), SP Sepharose FF (Pharmacia) (lane III) and affinity chromatography on amylose (New England Biolabs) (lane IV). The purified fusion protein was then cleaved with factor Xa (New England Biolabs) (lane V) and homogenous P1G-LlFpg was isolated from proteolysate by perfusion chromatography on HS POROS (PerSeptive Biosystem) (lane VI). Purification fractions (lanes I–VI) were analyzed by SDS–PAGE (Fig. 1). After mass spectroscopy and N-terminal micro-sequencing checking, the homogenous P1G-LlFpg (5–10 mg ml^{-1}) was stored at 193 K in 10 m*M* HEPES–NaOH pH 7.6, 100 m*M* NaCl, 5% glycerol, 5 m*M* β-mercaptoethanol and 0.1 m*M* phenylmethylsulfonyl fluoride.

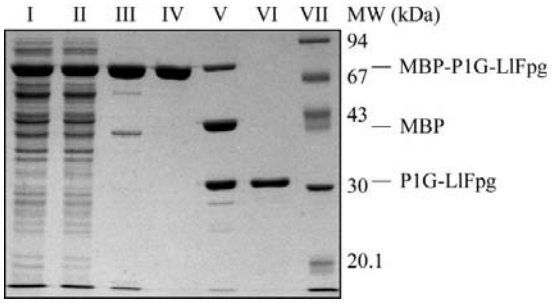

Figure 1
SDS–PAGE analysis of the P1G-LlFpg protein purification fractions. Lane VII, molecular mass markers (LMW, Pharmacia). Lanes I–VI correspond to the purification steps described in the text.

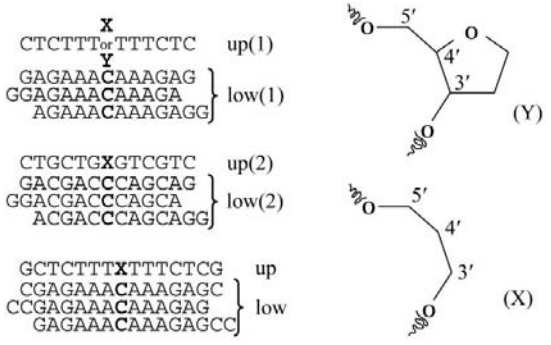

Figure 2
Oligonucleotides used for the crystallization assays. AP-site analogue containing strands (up) were annealed with their complementary strands (low) to generate 13-mer and 15-mer DNA duplexes (up/low) with blunt ends or one complementary 3′ or 5′ overhanging base ends. In all DNA duplexes, the 1,3-propanediol (X) or tetrahydrofuran (Y) site analogue was opposite the bold cytosine. (1) and (2) indicate the 13-mer sequences used in this study.

Self-complementary synthetic oligonucleotides (Fig. 2) were purchased from Eurogentec (Belgium). After purification by anion-exchange chromatography on MONO-Q (Pharmacia) and desalting by reverse-phase HPLC on ODS-Ultrasphere (Beckman), the oligonucleotides containing an AP-site analogue were annealed with their complementary strand to produce 12 DNA duplexes.

2.2. Systematic crystallization studies

In protein/DNA co-crystallization, the precise choice of DNA fragments is critical because interactions between DNA molecules are often responsible for intermolecular contacts in the crystal. Indeed, in many cases the DNA packs end-to-end, forming a pseudo-continuous helix (Anderson *et al.*, 1984; Jordan *et al.*, 1985). Thus, our strategy consisted of considering as crystallization variables the following DNA features: the length, sequence, recognition site and end nature of the DNA fragments (blunt-ended or sticky-ended). We used as DNA target sites two non-metabolizable substrate analogues, 1,3-propanediol and tetrahydrofuran AP-site derivatives, which are able to form very stable non-covalent complexes with Fpg (Castaing *et al.*, 1992, 1999). We then empirically selected several 13-mer and 15-mer double-stranded oligonucleotides with or without a complementary 3′ or 5′ overhanging base (Fig. 2). DNA duplexes were mixed with P1G-LlFpg and the 12 resulting complexes were tested for co-crystallization. Each protein–DNA complex solution was concentrated to a final concentration of 5 mg ml^{-1} with a 1.3-fold molar excess of DNA. Using the sparse-matrix crystallization screening kit (Hampton Research; Jancarik & Kim, 1991), crystallization was performed by the hanging-drop vapour-diffusion method. An initial systematic screen for crystallization was carried out at 293 K. Under several crystallization conditions and with various DNA/protein combinations, small crystals appeared within few days, whereas under the same conditions the protein alone did not crystallize. In all successful assays, the crystal solutions contained polyethylene glycol (20–30% PEG 4K or 8K) and salt (0.1–0.2 *M*) within a large pH range (Hampton Research Crystal Screen 1 solutions 6, 9, 10, 15, 17, 18, 22, 40, 42 for the complexes with 13-mer DNA duplexes and solutions 6, 9, 15, 17, 18, 20, 31, 38 for the complexes with 15-mer DNA duplexes). The larger crystals were obtained with complexes between P1G-LlFpg and the 1,3-propanediol AP-site analogue containing blunt-ended 13-mer (sequence 1) or one overhanging base ended 15-mer DNA. Interestingly, no Hampton Research Crystal Screen 1 solutions led to crystals with the complex containing the blunt-ended 15-mer DNA duplex. When the complexes containing sticky-ended DNA were used, crystallization occurred under different conditions according to whether the DNA ends were 5′ or 3′ overhanging. Thus, all these observations indicate clearly that the nature of the DNA ends plays a critical role in the crystallization process. This systematic crystallization screening also showed the effect of the nature of the

substrate analogue containing DNA [1,3-propanediol (X) or tetrahydrofuran (Y) in 13-mer sequence 1; Fig. 2]. Indeed, when the same DNA duplex contained the X or Y site, protein–DNA complexes with the blunt-ended X-DNA crystallized more easily than those with sticky-ended X-DNA and conversely for Y-DNA. These data suggest that the X-DNA and Y-DNA intrinsic structures and/or those induced by the Fpg binding are not equivalent and generate different crystal packings. This observation can also be correlated with the differential Fpg affinity for these substrate analogues (for LlFpg, K_a = 0.25 and 0.125 nM^{-1} for X and Y, respectively; Castaing *et al.*, 1999). Finally, the nature of the sequence (13-mer sequences 1 or 2, Fig. 2) appeared to have some effects on the crystallization, but these effects are less spectacular than those related to the other parameters.

The first generation of crystals was used for microseeding experiments to obtain single crystals with dimensions up to 0.1 mm (Fig. 3). The resulting crystals were then systematically tested at the BM30 beamline of the ESRF. The best conditions (0.1 *M* Tris–HCl pH 8.5, 0.2 *M* Li_2SO_4 or 0.1 *M* $MgCl_2$, 30% PEG 4K) were found for the complex between P1G-LlFpg and the blunt-ended 13-mer (sequence 1) DNA duplex containing the 1,3-propanediol AP-site analogue. These conditions gave similar crystals to the wild-type enzyme (data not shown). An electrophoresis mobility shift assay (Castaing *et al.*, 1999) of the content of carefully washed and dissolved diffracting crystals demonstrated the presence of the protein–DNA complex (data not shown).

2.3. Diffraction data collection

After optimization of the crystallization conditions, the best crystals, obtained in 28% PEG 4K, 0.2 *M* Li_2SO_4 and 0.1 *M* Tris–HCl pH 8.9, were soaked for several seconds in a solution containing 16% glycerol as a cryoprotectant and frozen in liquid nitrogen. Using synchrotron radiation at the BM30 beamline of the ESRF (Grenoble, France), we collected three data sets to 2.1 Å resolution at 100 K around the absorption edge of the intrinsic Fpg Zn atom. Diffraction data were processed with *DENZO* (Otwinowski & Minor, 1997) and scaled with *SCALEPACK* (Evans, 1993) (Table 1). Crystals belonged to the monoclinic system with space group $P2_1$ and unit-cell parameters a = 69.1, b = 62.6, c = 80.7 Å, β = 104.1°. Assuming two complex molecules to be present in the asymmetric unit, the solvent content was calculated to be about 46%.

Table 1
Native data-collection statistics.

Values in parentheses refer to the last resolution shell.

Wavelength (Å)	1.2831	1.2828	1.2714
Resolution (Å)	2.09	2.09	2.07
Total observations	103760	102393	93486
Unique reflections	38588	38044	37889
Data completeness (%)	97.6	96.6	93.5
Average redundancy	2.7	2.7	2.5
Mean $I/\sigma(I)$	14.0	9.7	11.4
R_{sym}†	0.033 (0.140)	0.035 (0.139)	0.048 (0.092)
R_{ano}‡	0.039	0.041	0.052

† R_{sym} is a measure of the internal consistency of the data defined as $\sum_h \sum_i I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i I_{h,i}$, where I_i is the intensity of the ith observation and $\langle I \rangle$ is the mean intensity of the reflection. ‡ R_{ano} is a measure of the mean relative anomalous difference between the Bijvöet pairs defined as $\sum |I(+) - I(-)| / \sum |\langle I(+) \rangle + \langle I(-) \rangle|$.

3. Conclusions

Thanks to a systematic crystallization study, we have obtained crystals of LlFpg bound to a DNA fragment containing a model of DNA lesion non-metabolizable by the enzyme. Despite the appearance of crystals in several conditions, only a few of them led to single crystals suitable for X-ray diffraction analysis. Therefore, of 12 DNA duplexes only one (13-mer, blunt ends, X site) led to crystals of diffraction quality. Thus, this systematic study also provides a good example of the striking effects of the nature of DNA on crystallization conditions of protein–DNA complexes.

We were able to collect high-resolution data using synchrotron radiation and are now involved in solving the crystal structure of the Fpg–DNA complex using the molecular-replacement method and/or MAD phasing.

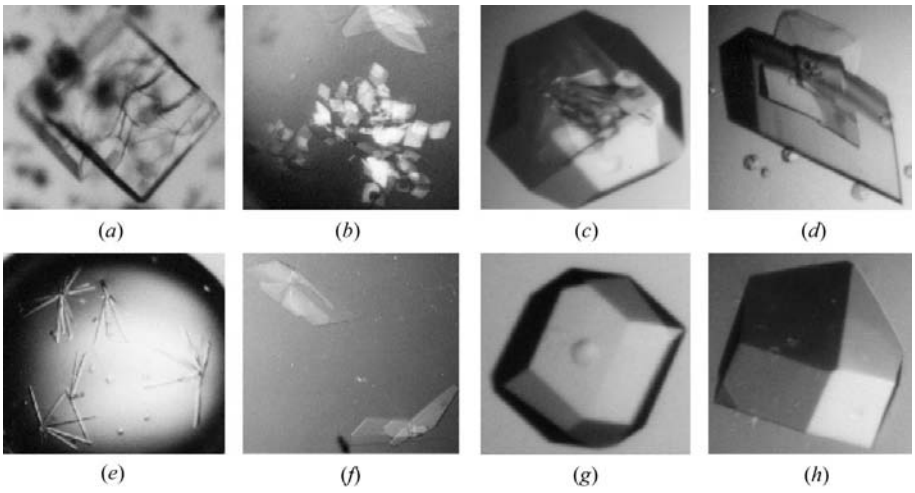

Figure 3
Typical P1G-LlFpg–DNA crystals obtained by a systematic screening. Under various crystallization conditions [20–30% PEG 4K and 0.1–0.3 *M* Li_2SO_4, 0.05–0.2 *M* $MgCl_2$ or 0.1–0.2 *M* $(NH_4)_2SO_4$], crystals grew at 293 K in the pH range 6–9. (*a*) Many thin plate blocks (0.5 × 0.5 mm), (*b*, *f*) diamond-shaped thin plates (0.1 × 0.1 mm), (*c*, *g*, *h*) large prism-shaped crystals, (*d*) very thick plates (0.6 × 0.3 × 0.15 mm), (*e*) needle-shaped crystals (0.2 mm). These crystals were obtained with the blunt-ended 13-mer duplex (sequence 1) (*b*, *d*, *e*, *f*) or with the one base overhanging 15-mer duplexes (crystals *a*, *c*, *g*, *h*), both containing the 1,3-propanediol abasic site analogue. Crystals of Fpg–15-mer diffracted to 5.0 Å and, after crystallization optimization, those obtained with 13-mer diffracted to 2.1 Å. Under the same conditions, crystals with similar faces were obtained from the wild-type enzyme.

We thank Nicole Bureaud for technical assistance. This work was supported by Electricité de France (EDF) and by the Comité de la Région Centre de la Ligue Nationale Contre le Cancer (Loiret, France).

References

Anderson, J., Ptashne, M. & Harrison, S. C. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 1307–1311.
Bailly, V., Verly, W. G., O'Connor, T. R. & Laval, J. (1989). *Biochem. J.* **262**, 585–589.
Boiteux, S., O'Connor, T. R. & Laval, J. (1987). *EMBO J.* **6**, 3177–3183.
Castaing, B., Boiteux, S. & Zelwer, C. (1992). *Nucleic Acids Res.* **20**, 389–394.
Castaing, B., Fourrey, J. L., Hervouet, N., Thomas, M., Boiteux, S. & Zelwer, C. (1999). *Nucleic Acids Res.* **27**, 608–615.
Castaing, B., Geiger, A., Seliger, H., Nehls, P., Laval, J., Zelwer, C. & Boiteux, S. (1993). *Nucleic Acids Res.* **21**, 2899–2905.
Chetsanga, C. J. & Lindahl, T. (1979). *Nucleic Acids Res.* **6**, 3673–3683.
Czeczot, H., Tudek, B., Lambert, B., Laval, J. & Boiteux, S. (1991). *J. Bacteriol.* **173**, 3419–3424.
Duwat, P., de Oliveira, R., Ehrlich, D. S. & Boiteux, S. (1995). *Microbiology*, **141**, 411–417.

Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.
Hatahet, Z., Kow, Y. W., Purmal, A. A., Cunningham, R. P. & Wallace, S. S. (1994). *J. Biol. Chem.* **269**, 18814–18820.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jordan, S. R., Whitcombe, T. V., Berg, J. M. & Pabo, C. O. (1985). *Science*, **230**, 1383–1385.
Lloyd, R. S. & Lin, S. (1993). *Nucleases*, Vol. 2, edited by S. Lin, R. S. Lloyd & R. J. Roberts, pp. 263–316. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
Michaels, M. L., Pham, L., Cruz, C. & Miller, J. H. (1991). *Nucleic Acids Res.* **16**, 3629–3632.
Mikawa, T., Kato, R., Sugahara, M. & Kuramitsu, S. (1998). *Nucleic Acids Res.* **26**, 903–910.
Murphy, T. M. & Gao, M. J. (1998). *Plant Physiol.* **118**, 1535.
O'Connor, T. R., Boiteux, S. & Laval, J. (1989). *Nucleic Acids Res.* **16**, 5879–5893.
O'Connor, T. R., Graves, R. J., de Murcia, G., Castaing, B. & Laval, J. (1993). *J. Biol. Chem.* **268**, 9063–9070.
O'Connor, T. R. & Laval, J. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 5222–5226.
Ohtsubo, T., Matsuda, O., Iba, K., Terashita, I., Sekiguchi, M. & Nakabeppu, Y. (1998). *Mol. Gen. Genet.* **259**, 577–590.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sidorkina, O. & Laval, J. (2000). *J. Biol. Chem.* **275**, 9924–9929.
Sugahara, M., Mikawa, T., Kuramitsu, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y. & Kuramitsu, S. (2000). *EMBO J.* **19**, 3857–3869.
Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P. & Nishimura, S. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 4690–4696.
Zharkov, D. O., Rieger, R. A., Iden, C. R. & Grollman, A. P. (1997). *J. Biol. Chem.* **272**, 5335–5341.