Acta Cryst. (1998). D**54**, 1002–1004

Crystallization and preliminary X-ray diffraction studies of DNA polymerase from the thermophilic archaeon *Sulfolobus solfataricus*

Vassilios Nastopoulos,[a,b] Francesca M. Pisani,[c] Carmelinda Savino,[a]* Luca Federici,[a] Mosè Rossi[c] and Demetrius Tsernoglou[a] at [a]*Istituto Pasteur-Fondazione Cenci Bolognetti, Department of Biochemical Sciences, Università di Roma 'La Sapienza', P. le A. Moro, 5, 00185 Rome, Italy,* [b]*Department of Chemistry, University of Patras, 26500 Patras, Greece, and* [c]*Istituto di Biochimica delle Proteine ed Enzimologia, CNR, Via Marconi, 10, 80125 Napoli, Italy. E-mail: savino@biosig.sci.uniroma1.it*

(*Received 2 October 1997; accepted 9 February 1998*)

Abstract

The thermophilic and thermostable family B DNA polymerase from the archaeon *Sulfolobus solfataricus* (M_r of about 100 kDa) has been crystallized by the hanging-drop vapour-diffusion method at 294 K using ammonium sulfate as precipitant. The crystals belong to the monoclinic space group $C2$ with cell dimensions $a = 187.4$, $b = 68.5$, $c = 125.8$ Å and $\beta = 107.8^\circ$ and diffract up to 2.7 Å resolution on a rotating-anode X-ray source. Native data have been collected at 100 K. A heavy-atom derivative search is in progress.

1. Introduction

DNA polymerases are usually multifunctional enzymes which possess, in addition to their synthetic function, one or two degradative activities (5′→3′ and/or 3′→5′ exonucleases). In all living organisms they are the key components of multi-enzymatic complexes responsible for the accurate duplication and maintenance of the genetic information since they are involved in DNA replication and repair (Kornberg & Baker, 1992). DNA polymerases have been classified into three major families (A, B and C) on the basis of their sequence similarity to *Escherichia coli* DNA polymerases I, II and III, respectively (Braithwaite & Ito, 1993). A fourth group (family X) includes eukaryotic DNA polymerase β which is involved in DNA repair reactions. Despite the apparent variety of types, it has been demonstrated that a certain number of amino-acid residues critical for the synthetic or for the degradative activity are highly conserved among DNA polymerases of different families (Joyce & Steitz, 1994). In addition, this functional similarity has been extended to other classes of template-directed nucleotide polymerases, such as reverse transcriptases, RNA replicases and DNA-dependent RNA polymerases (Arnold *et al.*, 1995). Indeed, this notion has been confirmed by analysis of the three-dimensional structure of various distantly related nucleotide polymerases, including the Klenow fragment (KF) of *E. coli* DNA polymerase I (Beese *et al.*, 1993) and DNA polymerase from *Thermus aquaticus*, both belonging to the family A (Kim *et al.*, 1995), rat DNA polymerase β (family X, Davies *et al.*, 1994), HIV-1 reverse transcriptase (Jacobo-Molina *et al.*, 1993) and bacteriophage T7 RNA polymerase (Sousa *et al.*, 1993). The similarity among the above structures is striking, particularly at the level of the polymerase domain, which in all cases possesses the shape of a right hand with 'fingers, thumb, palm' subdomains and contains a deep groove to accommodate the nucleic acid substrate. Recently, the crystal structure of a pol α DNA polymerase from bacterophage RB69 (which belongs to the family B) has been determined (Wang *et al.*, 1997). The polymerase catalytic domain is structurally similar to the palm domain of the polymerases of family A and X, but the finger and the thumb domain are unlike those of the other polymerase of known structure.

In this context, we have undertaken a study on family B DNA polymerase from the thermoacidophilic archaeon *Sulfolobus solfataricus* (DNA pol *Ss*). This enzyme, which is a monomer of about 100 kDa and possesses thermophilic and thermostable DNA polymerase and 3′→5′ exonuclease activities, was found by primary structure alignments to belong to family B of the DNA polymerases (Pisani *et al.*, 1992). This group is quite heterogenous including, in addition to DNA polymerase II from *E. coli*, eukaryotic cellular and viral DNA polymerases, replicases from the bacteriophages T4 and φ29, and several other DNA polymerases from Archaea, which are currently utilized in PCR (polytherase chain reaction) protocols by virtue of their high fidelity and thermal-stability features (Braithwaite & Ito, 1993). Recently, DNA pol *Ss* has been demonstrated by limited proteolysis studies to possess a modular organization of its associated catalytic activities. It is composed of two protease-resistant domains. The amino-terminal half of the protein molecule is responsible for the 3′→5′ exonuclease activity, whereas the carboxyl-terminal one retains the polymerization function (Pisani & Rossi, 1994; Pisani *et al.*, 1996).

2. Materials and methods

2.1. *Crystallization*

For the crystallization studies presented herein, we have utilized the recombinant DNA pol *Ss*, overproduced in *E. coli* under the control of a T7 promoter and extensively purified in a soluble and active form (Pisani, unpublished results). Crystallization assays were carried out by the conventional hanging-drop vapour-diffusion technique (McPherson, 1990). The protein solution for crystallization was made up of 20 m*M* Tris–HCl, pH = 8.0, 5 m*M* β-mercaptoethanol, 5 m*M* $MgCl_2$, 400 m*M* NaCl. The purity of the protein was >98%, as estimated by silver-staining, following SDS–PAGE (Pisani *et al.*, 1996). A wide range of crystallization conditions were initially tested. Single crystals were grown by mixing equal volumes (3 µl) of protein solution and reservoir solution consisting of 32–36% saturated ammonium sulfate in 100 m*M* Hepes buffer, pH = 7.8–8.2, 1.0 m*M* β-mercaptoethanol and 1.0 m*M* sodium azide. The drops were vapour equilibrated against 1.0 ml of the same reservoir solution at a temperature of 294 K. Although crystals grew from a range of protein concentrations (5.0–15.0 mg ml^{-1}), a concentration of 10 mg ml^{-1} gave less nucleation per drop as well as bigger and better diffracting

Table 1. *Statistical parameters of data collection on crystal of DNA pol Ss as a function of resolution*

Resolution (Å)	R_{merge}†	Multiplicity	Completeness (%)	$I > 3\sigma(I)$ (%)
20.0–6.23	0.040	2.8	91.5	95.2
6.23–4.61	0.065	2.8	94.2	90.2
4.61–3.82	0.082	2.7	95.0	85.9
3.82–3.34	0.167	2.7	96.2	70.9
3.34–3.00	0.366	2.6	96.1	46.5
Overall	0.101	2.7	94.3	72.9

† $R_{merge} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_h \sum_i I_h$, where I_{hi} is the ith observation of the reflection h, while $\langle I_h \rangle$ is the mean intensity of the ith reflection.

crystals. Small plate-like crystals appeared after approximately 4 months. They kept growing for the next 2–3 months and finally reached a maximum size of 0.9 × 0.4 × 0.2 mm. In their final shape some of the crystals are prism-like while others contain well defined curved surfaces. Analysis by SDS–PAGE of the protein from washed and dissolved crystals showed that what we have crystallized is indeed the DNA polymerase. Macroseeding techniques (Thaller *et al.*, 1981), as well as variation of the crystallization conditions (pH, precipitant concentration, addition of small amounts of organic solvents, glycerol) did not significantly improve the crystal size and diffraction quality of the crystals. It was also proven that the presence of Mg^{2+} is essential for crystallization.

2.2. *Cryofreezing and data collection*

Crystals were measured using a Rigaku R-Axis IIC imaging-plate detector, mounted on a Rigaku RU-200 rotating-anode X-ray (Cu $K\alpha$) generator, equipped with focusing mirrors and operating at 50 kV and 100 mA. The crystals are stable in the sealed well but they are not resistant in the X-ray beam at room temperature as they diffract only at lower resolution after few hours exposure in the X-ray beam. Because of the size of the unit cell and space group (see later), it is difficult to collect a complete data set at high resolution and at room temperature from a single crystal. However, it was possible to collect complete data from a single crystal by using low-temperature techniques (Rodgers, 1994). Cryofreezing was accomplished by the addition of glycerol. A crystal of DNA pol *Ss* was placed in 200 µl of mother liquor plus 50 µl of glycerol. A fiber loop (Hampton Research, Laguna Hills, California, USA) was used to fish the crystal out of cryosolvent and place it rapidly in the nitrogen gas coldstream maintained at 100 K, using an Oxford Cryosystems Cryostream. Diffraction spots can be observed up to 2.7 Å resolution (Fig. 1). Data at 3.0 Å were processed with *DENZO* (Otwinowski & Minor, 1996) and scaled with *SCALA* in the *CCP*4 package (Collaborative Computational Project, Number 4, 1994).

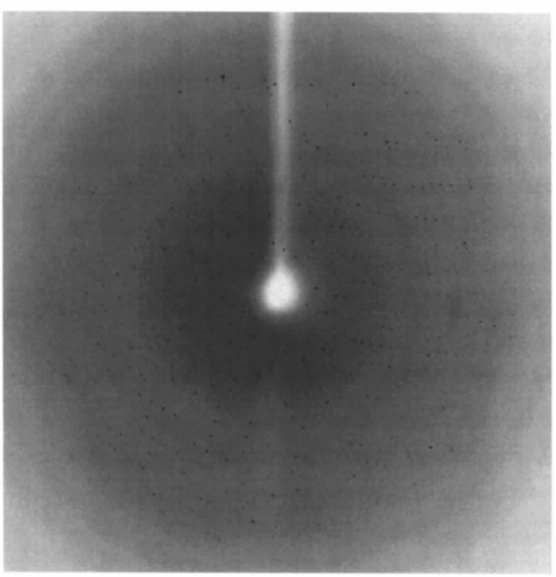

Fig. 1. A typical diffraction pattern of a crystal of DNA polymerase from *Sulfolobus solfataricus*. The edge of frame is at 2.6 Å.

3. Results and discussion

Diffraction is consistent with the space group $C2$. The room-temperature diffraction data show cell parameters of a = 189.1, b = 68.5, c = 130.0 Å and β = 110.0°. Data collected at 100 K had a smaller unit-cell volume with a = 187.4, b = 68.5, c = 125.8 Å and β = 107.8°. Diffraction spots were seen at 2.7 Å resolution, but the data set is complete to 94.3% at 3.0 Å, with an R_{merge} of 10.1% (based on intensities). The multiplicity is 2.7 and the percentage of data with $I > 3\sigma(I)$ is 72.9%. The completeness in the shell between 2.90 and 2.70 Å is less then 40%. Data-collection statistics are summarized in Table 1.

The molecular weight of the monomer is about 100 kDa, and assuming one molecule per asymmetric unit we obtain a crystal volume per protein mass (V_m) of 3.96 $Å^3$ Da^{-1}, corresponding to a solvent volume fraction of about 69%, which is rather high but within the range of V_m values for protein crystals (Matthews, 1968). An alternative possibility is that we have two molecules in the asymmetric unit. This leads to a V_m of 1.98 $Å^3$ Da^{-1} which is still acceptable. However, because of the rather large size of the polymerase we consider the latter possibility as the less likely. Screening of heavy-atom derivatives to solve the structure by the multiple isomorphous replacement method is under way and a number of these results appear very promising.

Special thanks are due to Professor M. Brunori for expert advice. We acknowledge FEBS for a short-term Fellowship (VN). This work was partially supported by the CNR (PS 'Biologia Strutturale') and by European Union Contract B104-CT96-0488. The X-ray diffraction facility was acquired with grants from MURST and the CNR.

References

Arnold, E., Ding, J., Hughes, S. H. & Hostomsky, Z. (1995). *Curr. Opin. Struct. Biol.* **5**, 27–38.

Beese, L. S., Derbyshire, V. & Steitz, T. A. (1993). *Science*, **206**, 352–355.

Braithwaite, D. K. & Ito, J. (1993). *Nucleic Acids Res.* **21**, 787–802.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davies, J. F., Almassy, R. J., Hostomska, Z., Ferre, R. A. & Hostomsky, Z. (1994). *Cell*, **76**, 1123–1133.
Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D. Jr, Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A., Clark, P., Hizi, A., Hughes, S. H. & Arnold, E. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 6320–6324.
Joyce, C. M. & Steitz, T. A. (1994). *Annu. Rev. Biochem.* **63**, 777–822.
Kim, Y., Eom, S. H., Wang, J., Lee, D.-S., Suh, S. W. & Steitz, T. A. (1995). *Nature (London)*, **376**, 612–616.
Kornberg, A. & Baker, T. A. (1992). *DNA Replication*, 2nd ed. New York: Freeman.
McPherson, A. (1990). *Eur. J. Biochem.* **189**, 1–23.
Matthews, B. M. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **234**, 129–132.
Pisani, F. M., De Martino, C. & Rossi, M. (1992). *Nucleic Acids Res.* **20**, 2711–2716.
Pisani, F. M., Manco, G., Carratore, V. & Rossi, M. (1996). *Biochemistry*, **35**, 9158–9166.
Pisani, F. M. & Rossi, M. (1994). *J. Biol. Chem.* **269**, 7881–7892.
Rodgers, D. W. (1994). *Structure*, **2**, 1135–1140.
Sousa, R., Chung, Y. J., Rose, J. P. & Wang, B. C. (1993). *Nature (London)*, **364**, 593–599.
Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R. & Jansonius, J. N. (1981). *J. Mol. Biol.* **147**, 465–469.
Wang, J., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., Konigsberg, W. H. & Steitz, T. A. (1997). *Cell*, **89**, 1087–1099.