Crystallization of the catalytic domain of murine terminal deoxynucleotidyl transferase

N. Sukumar, J. B. Boulé, N. Expert-Bezançon, N. Jourdan, J. Lescar, F. Rougeon, C. Papanicolaou and M. Delarue
Crystallization of the catalytic domain of murine terminal deoxynucleotidyl transferase

The catalytic domain of murine terminal deoxynucleotidyl transferase (TdT) has been crystallized in the space group P2₁2₁2₁, with unit-cell parameters a = 47.1, b = 86.2, c = 111.7 Å. The crystals diffract to a resolution of 2.4 Å using synchrotron radiation and a full data set has been collected from the native crystals. The enzyme was shown to be active in the crystalline state.

1. Introduction

Terminal deoxynucleotidyl transferase (TdT; E.C. 2.7.7.31) is a template-independent DNA polymerase which carries out the 3'-OH extension of a DNA strand one nucleotide at a time (Kato et al., 1967; Bollum, 1974, 1978). It cannot synthesize a strand de novo and the oligonucleotide primer must be at least three nucleotides long with a 5'-phosphate (Kato et al., 1967). TdT can incorporate both ribonucleotides and deoxyribonucleotides, as well as several unnatural nucleoside triphosphates. In vivo, it is known to contribute to the generation of diversity of the immune repertoire by adding random nucleotides, called N regions, at the V(D)J recombination junction sites of immunoglobulins and T-cell receptors (Kallenbach et al., 1990; Komori et al., 1993; Gilfillan et al., 1993).

TdT belongs to an ancient nucleotidyl-transferase (NT) superfamily, as shown by sequence database searches as well as by structure-inspired delineation of functionally important sequence motifs (Holm & Sander, 1995). This family includes DNA polymerase β (pol β), poly(A) polymerase, CCA:tRNA nucleotidyltransferase, 2′-5′ oligoadenylate synthetase and several other prokaryotic and eukaryotic nucleotidyltransferases (for a recent review, see Aravind & Koonin, 1999). Recent data indicate that Escherichia coli DNA polymerase III could also belong to the NT family (Pritchard & McHenry, 1999).

Only one member of this family, DNA polymerase β, is known in detail, at atomic resolution, both as an isolated protein and in various binary and ternary complexes with natural substrates (Davis et al., 1994; Pelletier et al., 1994a,b).

The topology of the catalytic core of pol β is very different from that of polymerases from the DNA pol I, reverse transcriptases and DNA polymerase α families. These three families share a common fold represented by the topology of the Klenow fragment (Ollis et al., 1985; Delarue et al., 1990; Wang et al., 1997; Doublié et al., 1998). Thermus aquaticus DNA polymerase I, a member of the pol I family, has been shown to adopt a closed and an open conformation during catalysis (Li et al., 1998). Pol β also exists in at least two conformations, namely the closed and open complexes (Sawaya et al., 1997). Since TdT does not use a template DNA strand, it would be interesting to know whether or not this transition between an open and closed complex also occurs.

Significant sequence similarity with pol β can be detected using BLASTP (E = 1 × 10⁻23 between rat pol β and murine TdT sequences). This similarity encompasses both the catalytic domain and the so-called finger and thumb regions, which are involved in the binding of both primer and template strands. Therefore, in principle TdT has all the elements to be a replicase. The structure should explain why TdT cannot accommodate a template strand and suggest possible mutations to transform it into a replicative polymerase.

The active site of all structurally known polymerases always contains crucial aspartate side chains in the vicinity of a catalytically important Mg²⁺ ion and of the incoming nucleoside triphosphate. In TdT, Co²⁺ and Mg²⁺ are the most efficient divalent cations for the incorporation of pyrimidine and purine nucleotides, respectively (Bollum, 1978). The three-dimensional structure of TdT should explain why TdT cannot accommodate a template strand and suggest possible mutations to transform it into a replicative polymerase.

In conclusion, structural comparison of polymerase members of the NT family should give insights into both the mechanisms of nucleic acid synthesis which are shared by all polymerases and those which underlie substrate specificities.

2. Materials and methods

2.1. Protein purification

The protein was overexpressed and purified in E. coli as described in Boulé et al. (1998).
The purity was estimated to be more than 98% based on SDS gel electrophoresis stained by Coomassie blue.

2.2. Crystallization

The protein was kept frozen at 253 K in 20 mM Tris–HCl pH 6.8, 200 mM NaCl, 50 mM Mg(OAc)₂, and 100 mM (NH₄)₂SO₄ at a concentration of 10 mg ml⁻¹. Aliquots of the protein stock solution were defrosted just before setting up the drops. Crystallization trials were performed using the hanging-drop method.

2.3. Crystal handling and flash-cooling

In a first step, crystals were isolated from their original drop using cryoloops from Hampton Research Corp. (Laguna Niguel, CA, USA) and were then transferred to 0.2 ml wells soaked in 400 mM glycerol as a cryoprotectant. Crystals of the TdT protein variants that were overexpressed in E. coli and purified, only the so-called Δ129 construct, in which the first 129 amino acids were removed, gave crystals diffracting to high resolution. It contains 399 amino acids (six histidines were added at the N-terminus to facilitate purification, as well as 14 other amino acids arising from the cloning itself) and is catalytically fully active (Boulé & Papanicolaou, data not shown).

Crystals of the Δ129 construct were obtained after less than 24 h of equilibration of the drop against a reservoir containing 20–23% PEG 6000, 100 mM MES pH 6.0, 1 M LiCl, 16 mM Mg²⁺, 16 μM Zn²⁺, 4 mM dATP and 400 mM [γ⁻³²P]-dATP. It was verified under the microscope that this treatment left the crystals intact. 200 mM EDTA was added to prevent any further reaction occurring before the crystals were dissolved in a formamide dye solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol). The mixture was then analysed using polyacrylamide gel electrophoresis.

2.4. TdT activity in the crystalline state

Carefully washed crystals were soaked for 12 h at 298 K in a solution containing 20% PEG 6000, 100 mM MES pH 6.0, 1 M LiCl, 16 mM Mg²⁺, 16 μM Zn²⁺, 4 mM dATP and 400 mM [γ⁻³²P]-dATP. It was verified under the microscope that this treatment left the crystals intact. 200 mM EDTA was added to prevent any further reaction occurring before the crystals were dissolved in a formamide dye solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol). The mixture was then analysed using polyacrylamide gel electrophoresis.

2.5. Data collection

All data collection was performed at the ESRF in Grenoble at beamlines ID2, ID14-EH3 or ID14-EH2. Diffraction data were collected from a crystal mounted in a cryoloop placed in a stream of cold nitrogen at 100–110 K. Exposure times were typically 3 × 5 s at beamline ID14-EH3 equipped with a MAR CCD detector. The wavelength was 0.9402 Å. The crystal-to-detector distance was set to 150 mm and the oscillation angle for each frame was 1°, collected in three passes. The speed of data collection using a CCD detector was instrumental in collecting a full data set at the highest possible resolution. A low-resolution pass at 3.5 Å was performed after the high-resolution pass, with exposure time 3 × 1 s. The crystals were stable in the beam for about 90–120 min. Data processing was performed with the programs DENZO and SCALEPACK (Otwinowski & Minor, 1997).

3. Results and discussion

TdT contains two functionally independent regions (Chang et al., 1982). The N-terminal region (130 amino acids) contains a BRCT motif involved in protein–protein interactions (Koonin et al., 1996; Caillabaut & Mornon, 1997; Bork et al., 1997). The BRCT domain of XRCC1 is believed to be functionally important in bringing different partners into physical proximity through heterodimerization (Rice, 1999). The 3.2 Å X-ray structure of this domain has recently been solved (Zhang et al., 1998). The C-terminal region of TdT contains the catalytic core of the protein.

The mass spectrometry of dissolved crystals revealed a major peak at m/z = 45505 ± 4.5, whereas the expected m/z is 45504.9 without the first (methionine) residue. N-terminal chemical sequencing was performed for the first 22 amino acids of the purified protein. They were found to be in accordance with the expected sequence, starting just after the N-terminus methionine. In addition to this major peak in the mass spectrometry experiment, an additional peak was observed at m/z = 45811 ± 1, which was consistently found in four separate experiments of different protein...
preparations; this was interpreted as an adduct between a cysteine residue of the protein and a glutathione molecule.

To assess the presence of TdT activity in the crystals, crystals were first washed several times in the reservoir solution and then incubated overnight in a solution containing a radioactive \((dA)_10\) oligonucleotide primer and dATP. PAGE analysis of the reaction products clearly showed primers elongated by several nucleotides. This result suggests that the enzyme is active in the crystal form (Fig. 2).

Using synchrotron radiation, a full data set was collected from a single native crystal at beamline ID2 at 2.8 Å resolution; subsequently, a full data set at 2.4 Å was collected at beamline ID14-EH3. A total of 110 frames were collected from a single crystal.

The unit-cell parameters are \(a = 47.1, b = 86.2, c = 111.7\) Å and the space group is \(P2_12_12_1\). An overall total of 99 454 observations leading to 19 351 independent reflections were recorded (99.8% complete for all data and 98.8% complete for the last resolution shell), with an average \(R_{sym}\) of 5.9% (23% in the last resolution shell, 2.46–2.40 Å resolution). The percentage of data with \(I > 3\sigma(I)\) is 84 and 61.2% for all data and for the last resolution shell, respectively. The mean redundancy is 5.2 for all reflections and 4.1 for the last resolution shell.

Molecular replacement using DNA polymerase \(\beta\) has so far been unsuccessful. This may be because of the relatively low sequence identity between the two sequences (27.5% for 332 aligned positions).

Heavy-atom derivatives are currently being screened in order to solve the phase problem experimentally.

We thank A. Namane for performing the mass-spectrometry measurements and J. D’Alayer for chemical N-terminal sequencing. It is a pleasure to thank S. Arzt and W. Burmeister for help during data collection at beamline ID14-EH3, as well as E. Mitchell at beamline ID14-EH2, at the ESRF in Grenoble, France. This work was partly supported by a grant from the Association de la Recherche sur le Cancer (ARC). NJ was supported in part by a fellowship from the Institut Pasteur (Bourse Cantarini). NS was also supported by the latter fellowship.

References


