Crystallographic studies of *Thermus aquaticus* RNA polymerase

S.A. Darst, G. Zhang, E.A. Campbell, and C. Richter
The Rockefeller University, 1230 York Avenue, New York, NY 10021 USA

Introduction

Transcription is the major control point of gene expression and RNA polymerase (RNAP) is the central enzyme of transcription. The RNAP in all cellular organisms is a complex molecular machine. In its simplest bacterial form, the enzyme comprises at least four subunits with a total molecular mass of about 400 kDa. The eukaryotic enzymes comprise upwards of a dozen subunits with a total molecular mass of about 500 kDa. The essential core component of the bacterial RNAP (subunits $\alpha_2\beta\beta'$) is evolutionarily conserved from bacteria to man, making the bacterial RNAPs excellent model systems.

To provide a more detailed framework for interpreting the existing genetic, biochemical, and biophysical information, as well as to provide structures with more predictive power to guide further studies aimed at understanding the transcription process and its regulation, three-dimensional crystals suitable for x-ray crystallography are required. Despite extensive effort, suitable crystals of the *Escherichia coli* enzyme, which is by far the best characterized, were never obtained. Switching to the study of RNAP from the thermophilic eubacteria *Thermus aquaticus* (*taq*) provided the breakthrough that was needed [1].

Methods and Materials

Methods and materials are described in [1].

Results

Tetragonal crystals of the *taq* core RNAP (subunit composition $\alpha_2\beta\beta'\omega$), space group $P4_12_12$ ($a = b = 201$, $c = 294$ Å), were grown by vapor diffusion. The crystals contained one 375.4 kDa core RNAP molecule per asymmetric unit. The frozen crystals diffracted to 4.0 Å in house. Spots could sometimes be observed to 2.7 Å resolution at synchrotron beamlines (Cornell High Energy Synchrotron Source A1/F1, National Synchrotron Light Source X25, and Advanced Photon Source BioCARS). Useful diffraction from the radiation-sensitive crystals was anisotropic, with reflections along the best and worst directions at 3.0 Å and 3.4 Å, respectively. Longer exposures or larger crystals did not improve the resolution limit.

The structure was solved by multiple isomorphous replacement (MIR), using a combination of metal clusters and conventional derivatives, in part using data collected at BioCARS [1]. The initial MIR map showed protein-solvent boundaries and contained some identifiable α-helices. Density modification using SOLOMON resulted in a dramatically improved map. The fold of the *E. coli* α-subunit N-terminal domain (NTD) dimer, previously solved in our laboratory [2], was easily recognized, and the αNTD structure was modeled as described [1]. Phase combination and multidomain noncrystallographic symmetry averaging were then used to obtain a slightly improved map. This map was exceptionally clean, with secondary structural elements and well-connected main-chain density over most of the structure, allowing building of a polyalanine model containing about 85% of the expected number of residues for β and β'.

Side-chain density, while present in much of the map, was weak to nonexistent in other regions. For this reason, selenomethionyl core RNAP was prepared and crystallized [1]. The resulting Fourier-difference peaks aided in the localization of methionine residues during modeling.
The current model contains about 70% of the main chain of β', the complete main chain of β (except for a few residues at each terminus), the αNTD dimer, a 91-residue polyAla model of ω, one Mg^{2+}-ion (chelated at the active center), and one Zn^{2+} ion. Lacking electron density and presumably disordered in the crystal are both α C-terminal domains, as well as a 74-residue segment of β' that includes a Zn^{2+}-binding motif in co-served region A. A nonconserved sequence of 330 residues in β' is also not modeled. Several stretches of residues are modeled as polyAla. The model is not yet well refined; the *R* factor is 0.33 for data from 8–3.3 Å resolution ($R_{free} = 0.39$). Nevertheless, there is an amazing congruence between the structure and a wealth of functional data on the enzyme [1]. Further refinement is in progress.

Discussion

In addition to structural and functional analysis of the enzyme guided by the 3.3 Å resolution structure, our current efforts include preparing and crystallizing the RNAP holoenzyme (containing the promoter-specificity subunit), crystallizing the core RNAP in complex with DNA (binary complex), and crystallizing an active transcription complex containing DNA and RNA (ternary complex).

Acknowledgments

Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38. G.Z. was supported by a Postdoctoral Fellowship from the National Institutes of Health. E.A.C. was supported by a Kluge Postdoctoral Fellowship.

References

[1] G. Zhang, E.A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S.A. Darst *et al*., *Cell* **98**, 811–824 (1999).
[2] G. Zhang and S.A. Darst, *Science* **281**, 262–266 (1998).