Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of the Oct-1/SNAP190 peptide/DNA complex

Stacy Hovde, Aimee Brooks, Katie Strong and James H. Geiger*

Michigan State University Chemistry Department, East Lansing, MI 48824, USA

Correspondence e-mail: geiger@cem.msu.edu

Received 14 September 2001
Accepted 13 December 2001

Crystals of the Oct-1 POU/SNAP190 peptide/DNA tertiary complex have been obtained by hanging-drop vapor diffusion at 293 K in 20% 2-propanol, 20% PEG 4000 and 0.1 *M* sodium citrate pH 5.6. The Oct-1 POU protein has two domains, one a homeodomain and the other a POU domain, which are connected by a flexible linker. The DNA used in the complex is slightly different in the octamer region compared with the two previously crystallized Oct-1 POU/DNA complexes. The DNA is 14 base pairs, with an octamer sequence of 5′-ATGTAGAT-3′ and an overhang of one base on both strands. The SNAP190 peptide is 27 amino acids long (residues 884–910). The crystals diffract to 2.3 Å (94.1% completeness) at the synchrotron under cryogenic (123 K) conditions. The crystals are triclinic, space group *P*1, with unit-cell parameters $a = 36.4$, $b = 54.9$, $c = 77.6$ Å, $\alpha = 94.9$, $\beta = 99.6$, $\gamma = 109.2^{\circ}$. This structure will provide insight into how Oct-1 interacts with SNAP190, a critical component of the small nuclear RNA-activating protein complex (SNAPc). Transcription of human snRNA genes is activated by these direct protein–protein interactions.

1. Introduction

Transcriptional regulation involves coordinated communication between transcriptional regulatory proteins and the general transcription machinery. We are studying the transcription of the human snRNA gene family in order to understand the mechanisms of transcription activation. These genes are amongst the most abundantly transcribed genes in the cell and, interestingly, transcription occurs by either RNA polymerase (Pol) II or III, depending upon promoter architecture. Promoters that are recognized by Pol III have both a TATA box and a proximal sequence element (PSE), while promoters transcribed by Pol II have only a PSE (Kunkel & Pederson, 1988; Lobo & Hernandez, 1989; Lobo *et al.*, 1990). Nevertheless, TATA-binding protein (TBP) is required for transcription of both of these classes of genes. Though all yeast Pol III genes and most Pol III genes from higher eukaryotes use some form of TFIIIB-dependent transcriptional initiation mechanism, the small nuclear RNAs of higher eukaryotes utilize the SNAPc complex to recruit the polymerase. The TATA-binding protein (TBP) is also recruited to the promoter through interactions with SNAPc (Sadowski *et al.*, 1993).

The five known components of the SNAPc complex vary in regard to size and function. The largest subunit, SNAP190, contains a small conserved region (residues 869–912) that interacts with Oct-1 and activates transcription. This region is homologous to the first 63 amino acids of OBF-1, a B-cell specific co-activator that associates with Oct-1 bound to octamer motifs and increases transcription from immunoglobulin promoters (Ford *et al.*, 1998). An apparently direct interaction between a basal initiation factor (SNAPc) and an activator (Oct-1) bypasses the need for a coactivator (OBF-1) in some contexts. We are interested in SNAPc because SNAPc-mediated transcription represents the only known TFIIIB-independent mechanism of Pol III initiation. In order to understand the interaction between SNAP190 and Oct-1, we have crystallized the complex on a U1 octamer site and are in the process of solving the three-dimensional structure.

The Oct-1 POU protein has been previously crystallized in complex with DNA (Klemm *et al.*, 1994) and in complex with DNA/OBF-1 (Chasman *et al.*, 1999). Both structures are at lower resolutions (3.0 and 3.2 Å, respectively) than our data for this complex (2.3 Å). These structures should help in obtaining a molecular-replacement solution. In addition, details of the linker and hydration information may be gleaned from this extension in resolution. We can also compare and contrast the interactions we observe with those of Oct-1 with OBF-1 (Bob-1, Oka-B). There are

important sequence differences between SNAP190 and OBF-1 in this interaction region.

2. Protein, DNA and SNAP190 peptide purification

The Oct-1 POU domain (residues 284–439) was prepared in *Escherichia coli* as a glutathione S-transferase (GST) fusion protein. The protein was purified on glutathione-agarose beads (Sigma) and cleaved with thrombin (Sigma) as described elsewhere (Aurora & Herr, 1992). The DNA was ordered from the Keck Oligonucleotide Synthesis Facility at Yale University and was purified as described previously (DeWees & Geiger, 1999). The SNAP190 peptide was also ordered from the Keck Yale facility. It was purified on a C18 column (Vydax) with an acetonitrile gradient. The peptide was purged with nitrogen and then lyophilized for use in complex formation. The Oct-1, DNA and SNAP190 were combined in a 1:1.2:3.0 molar ratio. The concentration of the Oct-1 protein was measured using the Bradford assay. The DNA concentration was calculated from its absorption at 260 nm. The peptide concentration was estimated using SDS–PAGE. The complex was then concentrated in a Centricon-3 (Amicon) so that the final Oct-1 POU concentration was around 5 mg ml^{-1}. The complex was exchanged into 10 m*M* Na HEPES pH 7.9, 10 m*M* DTT buffer by two cycles of concentration and dilution. Crystals did not form in the absence of DTT.

The complex was screened for crystallization in 2 µl hanging drops in a variety of crystallization solutions. The 2 µl hanging drop was equilibrated against 300 µl of crystallization solution in a 1:1 ratio. The 50-condition sparse-matrix crystal screen (Jancarik & Kim, 1991) yielded diffraction-quality crystals in 20% 2-propanol, 20% PEG 4000 and 0.1 *M* sodium citrate pH 5.6. The DNA used in the crystallization was one of several tried based on octamer motifs found in vertebrate snRNA genes (Hernandez, 1992). Fig. 1(*a*) depicts the DNA and sequence of the SNAP190 peptide used in crystallization. The crystals grew in a week to maximum dimensions of 700 × 50 × 25 µm (Fig. 1*b*) at room temperature.

3. Data collection and evaluation

The crystals were transferred to a cryoprotectant solution consisting of 30% MPD, 20% 2-propanol, 20% PEG 4000 and 0.1 *M* sodium citrate pH 5.6 and quick-frozen in liquid nitrogen. X-ray diffraction data to a resolution of 2.3 Å were collected at the Structural Biology Center ID19 beamline at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL, USA). Data were collected using a custom-built 3 × 3 array (3072 × 3072 pixels) CCD area detector. The crystal-to-detector distance was 200 mm and two data sets were collected on this crystal. The two were scaled together and a total of 430° of data were collected. Diffraction data were indexed and integrated using *HKL*2000 and *SCALEPACK* (Otwinowski & Minor, 1997).

The crystals are triclinic (space group *P*1), with unit-cell parameters a = 36.4, b = 54.9, c = 77.6 Å, α = 94.9, β = 99.6, γ = 109.2°. There are two molecules in the asymmetric unit, with a solvent content of 52%. X-ray diffraction data were 94.1% complete, with an $R_{\rm merge}$ of 0.096 for 24 916 unique reflections from a total of 110 313 measured reflections. Detailed data-collection statistics are given in Table 1. The structure was solved by molecular replacement using *AMoRe* (Navaza, 1994). The Oct-1 POU structure (Klemm *et al.*, 1994) was used as a model with only the few central base pairs of DNA included in the initial search model. The second molecule was found using only the POU domain as a search model. The homeodomain and the DNA portion of the second molecule were built using the first molecule as a guide. After an initial cycle of refinement, an $R_{\rm free}$ of 42% and an R factor of 33% were obtained excluding both peptides and most of the DNA. Clear electron density for both peptides was observed in this map, confirming the presence of the peptide in both molecules.

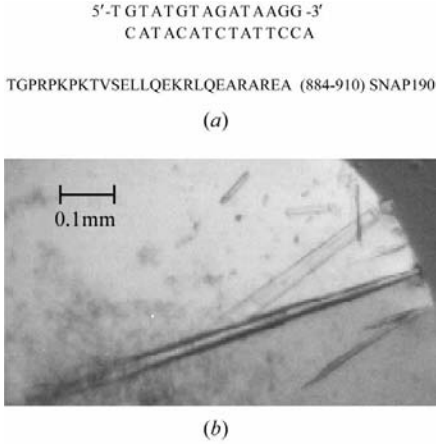

Figure 1
(*a*) U1 octamer DNA sequence used in crystallization of the complex. (*b*) Typical crystals of the complex, with dimensions of 0.7 × 0.05 × 0.025 mm.

Table 1
Data-collection statistics.

Values in parentheses refer to the last shell. Data was collected at the Advanced Photon Source, Structural Biology Center ID19 beamline.

Wavelength (Å)	0.979
Resolution range (Å)	50.0–2.3 (2.38–2.30)
Reflections measured	110313
Completeness (%)	94.1 (92.3)
$I/\sigma(I)$	12.7 (3.0)
$R_{\rm merge}(I)$† (%)	9.6 (36.9)
Overall χ^2 for scaling	0.748

† $R_{\rm merge} = \sum_I |I_I - \langle I\rangle|/\sum\langle I\rangle$, where I_I is an individual intensity measurement and $\langle I\rangle$ is the average intensity for this reflection, with summation over all data.

We would like to thank Dr R. William Henry and his laboratory for collaborating with us on this project, and Soren Ottosen for help with improving the yield of Oct-1 POU protein. The US Department of Energy, Basic Energy Sciences and Office of Science supported our use of the Advanced Photon Source. The authors also thank Rongguang Zhang and Andrzej Joachimiak from the Structural Biology Center at APS for their assistance throughout the data collection.

References

Aurora, R. & Herr, W. (1992). *Mol. Cell Biol.* **12**, 455–467.
Chasman, D., Cepek, K., Sharp, P. A. & Pabo, C. O. (1999). *Genes Dev.* **13**, 2650–2657.
DeWees, S. & Geiger, J. H. (1999). *Acta Cryst.* D**55**, 2039–2040.
Ford, E., Strubin, M. & Hernandez, N. (1998). *Genes Dev.* **12**, 3528–3540.
Hernandez, N. (1992). *Transcriptional Regulation*, edited by S. L. McKnight & K. R. Yamamoto, pp. 281–313. New York: Cold Spring Harbor Laboratory Press.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Klemm, J. D., Rould, M. A., Aurora, R., Herr, W. & Pabo, C. O. (1994). *Cell*, **77**, 21–32.
Kunkel, G. R. & Pederson, T. (1988). *Genes Dev.* **2**, 196–204.
Lobo, S. M. & Hernandez, N. (1989). *Cell*, **58**, 55–67.
Lobo, S. M., Ifill, S. & Hernandez, N. (1990). *Nucleic Acids Res.* **18**, 2891–2899.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–325.
Sadowski, C. L., Henry, R. W., Lobo, S. M. & Hernandez, N. (1993). *Genes Dev.* **7**, 1535–1548.