03-Crystallography of Biological Macromolecules

PS–03.01.05 Crystallization of Poliovirus RNA-dependent RNA Polymerase

Jun Tsao Center for Macromolecular Crystallography, University of Alabama at Birmingham, Al 35294

The single-stranded RNA genome of poliovirus is replicated by a virus encoded RNA-dependent RNA polymerase designated $3D^{pol}$. This enzyme is a monomer protein that has a molecular weight of 52 KD. The polymerase is essential for replication of the viral genome, interacting with the 3' end of both plus and minus-strand RNA, as well as various host proteins and other viral proteins. Limited mutagensis of the protein has indicated the importance of its linear sequence with relationship to function; site-directed point-mutants and linker-insertion mutants at most places in the polymerase gene lead to diminished, or in many cases, complete disruption of function. To further identify structure-function relationships of this protein, the determination of a high resolution structure is necessary. In addition to providing a better understanding of the structural basis for poliovirus RNA polymerase activity, this will also lead to informed targeted drug-design against the polio RNA polymerase, as well as many other related polymerases.

The $3D^{pol}$ gene encoding the polymerase has been cloned into an expression vector based on the T7 promoter (pET, Novagen). Expression of the enzymatically active $3D^{pol}$ is obtained by transformation of the plasmid into an *E.coli* strain containing the T7 polymerase gene on the host chromosome under control of the *lac* promoter (B21/DE3). The products of $3D^{pol}$ form inclusion bodies with an *E.coli* protein (~40 KD) in the cells. The first step of protein purification is to isolate the inclusion bodies. Then, urea is used to completely denature the inclusion bodies. The denatured $3D^{pol}$ is slowly re-natured upon dialysis. Further purification is acheived through a phosphate-cellulose column.

The crystallization of the purified $3D^{pol}$ is carried out by the hanging drop vapor diffusion method. Preliminary results have shown the formation of small crystals. The crystals were formed in approximately one month at room temperature. They have a hexagoinal morphology with a maximun size of 0.1 mm. The crystals diffracted X-rays to 3.5 Å at synchrotron stations.

Current experiments are emphasized at the optimization of conditions for obtaining large enough crystals for the determination of the three-dimensional structure of polivirus RNA polymerase.

PS–03.01.06 CRYSTAL STRUCTURE OF *Hha*I DNA METHYLTRANSFERASE COMPLEXED WITH S-ADENOSYL-L-METHIONINE

Xiaodong Cheng*, Sanjay Kumar+, Mi Sha, Richard J. Roberts+

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 11724. +present address: New England Biolabs, 32 Tozer Road, Beverly, MA 01915.

DNA methyltransferases are found in organisms ranging from bacteria to mammals. The DNA methyltransferase from the bacterium *Haemophilus haemolyticus* catalyzes the transfer of a methyl group from S-adenosyl-L-methionine (AdoMet) to C-5 of the internal cytosine in the DNA sequence GCGC. The three dimensional structure of the M.*Hha*I-AdoMet complex has been determined and refined at a resolution of 2.5 Å. The structure is the first to be solved for any DNA methyltransferase as well as being the first for any methyltransferase that utilizes the ubiquitous methyl donor AdoMet. Due to the conserved nature of (cytosine-5)-methyltransferases, the information obtained from this structure can be generalized to the entire family, including the mammalian CpG methyltransferase.

Crystals of M.*Hha*I-AdoMet complex were monoclinic space group $P2_1$ with unit-cell dimensions of a=55.3 Å, b=72.7 Å, c=91.0 Å, and ß=102.5°. The structure was solved with phases derived from two mercurial derivatives with anomalous scattering. The current crystallographic R-factor is 22.7% for 20057 reflections between 8.0 and 2.5 Å resolution with $F > 2\sigma(F)$.

The molecule of size 40 x 50 x 60 Å is folded into two domains connected by a hinge region. The large domain is a mixed α/ß structure consisting of the N-terminal two-thirds of the protein. It contains most of the invariant sequences amongst (cytosine-5) methyltransferases, and is responsible for providing the catalytic nucleophile Cys81 and the binding of cofactor AdoMet. The core of the large domain is a six-strand ß-sheet which contains four adjacent parallel strands and a hairpin. The small domain, predominated by ß-strands, contains the 'variable region' amongst different (cytosine-5)-methyltransferases, and is responsible for the DNA sequence specificity. Five up-and-down antiparallel strands are arranged in circular formation like the blades of a propeller. A hinge region, built up from an α-ß-α structure, connects the two domains. A cleft between the two domains is the likely active site, where the binding of AdoMet and DNA can occur in close proximity to the catalytic center. At the surface of the cleft, a cavity embedded in the large domain next to the carboxyl ends of the parallel ß-strands contains the AdoMet binding site. This cavity is adjacent to the active site nucleophile Cys81.

Mutagenesis of critical residues in *Hha*I methylase is underway. Replacement of the invariant cysteine (Cys81) with serine, arginine, histidine, and glycine abolished catalytic activity. In addition, the glycine substitution was cytotoxic to *E. coli* due to an extremely high affinity (about 100 fold higher than wild type) for DNA containing the recognition sequence.

