Selenium-assisted Nucleic Acid Crystallography: Use of DNA Phosphoroselenoates for MAD Phasing

C.J. Wilds,[1] R. Pattanayek,[1] C. Pan,[1] Z. Wawrzak,[2] M. Egli[1]
[1]Department of Biochemistry, Vanderbilt University, Nashville, TN, U.S.A.
[2]Advanced Photon Source (APS), Argonne National Laboratory, Argonne, IL, U.S.A.

Introduction

The combination of synchrotron radiation and a variety of atoms or ions (either covalently attached to the biomolecule prior to crystallization or soaked into crystals) that serve as anomalous scatterers constitutes a powerful tool in the x-ray crystallographer's repertoire of structure determination techniques. Phosphoroselenoates in which one of the nonbridging phosphate oxygens in the backbone is replaced by selenium offer a simplified means for introducing an anomalous scatterer into oligonucleotides by conventional solid-phase synthesis. Unlike the situation for other methods that are used to derivatize DNA or RNA by covalent attachment of a heavy atom (i.e., bromine at the C5 position of pyrimidines), tedious synthesis of specialized nucleosides is not required. Introduction of selenium is readily accomplished in solid-phase oligonucleotide synthesis by replacing the standard oxidation agent with a solution of potassium selenocyanide. This results in a diastereomeric mixture of phosphoroselenoates that can be separated by strong anion-exchange high-pressure liquid chromatography (HPLC). As a test case, all 10 DNA hexamers of the sequence CGCGCG containing a single phosphoroselenoate linkage (PSe) were prepared. Crystals were grown for a subset of them, and the structure of $[d(C_{PSe}GCGCG)]_2$ was determined by the multiwavelength anomalous dispersion (MAD) technique and refined to 1.1 Å resolution.

Methods and Materials

Oligonucleotide Synthesis and Purification

Oligonucleotides were synthesized on an Applied Biosystems Inc. (381A) DNA synthesizer following slight modifications to published procedures. All monomer coupling times were 90 seconds. In order to introduce the PSe linkage, solid-phase synthesis was halted after coupling of the phosphoramidite to the growing chain. Because the DNA synthesizer contains only one port for oxidizing agent, the column was removed from the synthesizer, and a saturated solution of KSeCN in 95% acetonitrile/5% triethylamine was added to the column and allowed to react for 24 hours in the dark, similar to the procedure carried out by Stein and coworkers to make a uniformly substituted PSe oligodeoxynucleotide via H-phosphonates. After that time had passed, the selenizing reagent was removed, and the controlled pore glass (CPG) support was washed with acetonitile. Capping was performed manually, and the column was returned to the synthesizer to continue extension of the sequence. All oligomers were synthesized with the 5′-terminal trityl group "off" and deprotected with concentrated ammonium hydroxide at 65°C for 8 hours. The sequences were analyzed and purified by strong anion exchange (SAX) HPLC by using a DNAPAC PA-100 analytical column (4 × 25 mm) purchased from Dionex Corporation (Sunnyvale, CA). For preparatory runs, 10 o.d. units at a time were purified on an analytical column using a gradient of 25 mM TrisHCl (pH = 7.8) to 0.5 M NaCl over 45 minutes at a flow rate of 1.0 mL/min. Oligomers purified by HPLC were desalted on Sep-Pak cartridges (Waters Corporation, Milford, MA). The cartridge was pre-equilibrated with acetonitrile followed by water. The oligonucleotide solution was applied to the cartridge and washed with water. A solution of 75% methanol in water was then used to elute the desalted oligomer. Molecular weights of the sequences were determined by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry.

Crystallization and Data Collection

Crystals of diastereomerically pure hexamers CGCGCG with single PSe linkages were grown by using the sitting-drop vapor-diffusion method. Crystallization conditions consisted of 5 mM DNA (2 μL), 200 mM sodium cacodylate (pH = 7, 5 μL), 15.6-62.5 mM $MgCl_2$ (5 μL), and 7.8-31.2 mM spermine tetrahydrochloride (5 μL) against a reservoir of 2-methyl-2,4-pentanediol (MPD, 30% in water, 25 mL), set up at room temperature. Diffraction-quality crystals typically appeared in 1 week. For data collection, $[d(C_{PSe}GCGCG)]_2$ crystals were picked up from a droplet with a nylon loop and transferred into a cold N_2 stream (120K). Crystals belonged to the orthorhomblic space group $P2_12_12_1$ with unit cell constants of a = 17.778 Å, b = 31.348 Å, and c = 44.116 Å. The precise locations of the inflection point and peak for the selenium absorption edge were determined from x-ray fluorescence spectra. A

commonly used buffer for crystallization of nucleic acids is cacodylate. The resulting high concentration of arsenic in the mother liquor and crystal resulted in a prominent peak in the x-ray fluorescence spectrum that should not be confused with the Se K edge (As at 11.8667 keV versus Se at 12.6578 keV). Data at three wavelengths from a single crystal were collected on DND-CAT insertion device beamline 5-ID at the APS. All data to a maximum resolution of 1.1 Å were integrated and scaled with the programs DENZO and SCALEPACK, respectively.

Structure Determination and Refinement

Selenium positions were determined with CNS-solve. The initial phases were improved with density modification, and the computed Fourier electron density maps were visualized with the program TURBO. The complete Z-DNA hexamer duplex was resolved in the maps, along with many water molecules. Dimer steps $[d(C_pG)]_2$ were taken from the crystal structure of the Z-DNA duplex with nucleic acid database code ZDF001 and built into the map. The initial refinement was carried out with the program CNS, 5% of the reflections set aside for calculating the R-free. The program SHELX-97 was used for the anisotropic refinement of all atoms and calculation of hydrogen atom positions. Along with the hexamer duplex, 73 water molecules, a magnesium hexahydrate ion, and a spermine molecule were included in the refinement that converged at values for R-work and R-free of 9.7% and 12.9%, respectively, based on all reflections to 1.1-Å resolution. Residues of one strand are numbered 1 through 6, and residues of the second strand are numbered 7 through 12.

Acknowledgments

This research is supported by the National Institutes of Health (Grant No. GM 55237 to M. Egli) and by the Petroleum Research Fund administered by the American Chemical Society (Grant No. 34538-AC4 to M. Egli). C.J. Wilds is the recipient of postdoctoral scholarships from the Natural Sciences and Engineering Research Council of Canada and the Fonds pour la Formation de Chercheurs et l'Aide a la Recherche. We thank Z. Huang and N. Carrasco, Brooklyn College, The City University of New York, for helpful discussions, and we are grateful to A. Noronha and P. Miller of Johns Hopkins University for running MALDI-TOF spectra. Portions of this work were performed at the DND-CAT beamline at sector 5 of the APS. DND-CAT is supported by E.I. DuPont de Nemours & Co., The Dow Chemical Company, the National Science Foundation through Grant No. DMR-9304725, and the State of Illinois through the U.S. Department of Commerce and Illinois Board of Higher Education, Higher Education Cooperation Act Grant IBHE HECA NWU 96. Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-102-ENG-38. This report was taken from C.J. Wilds, R. Pattanayek, C. Pan, Z. Wawrzak, and M. Egli, "Selenium-assisted nucleic acid crystallography: Use of DNA phosphoroselenoates for MAD phasing," J. Am. Chem. Soc. **124**, 14910-14916 (2002).