Some aspects of crystal centering during X-ray high-throughput protein crystallography experiment

Yu.A.Gaponov [a,*], N.Matsugaki [b], K.Sasajima [b], N.Igarashi [b], S.Wakatsuki [c]

aDepartment of atmospheric research, Novosibirsk State University, Pirogova 2, Novosibirsk, 630090, Russia
bStructural Biology Research Center, High Energy Accelerator Research Organization, 1-1 Oho, Tsukuba, 305-0801, Japan
cStanford University, Stanford, California, 94305, USA

Abstract

A set of algorithms and procedures of a crystal loop centering during X-ray high-throughput protein crystallography experiment has been designed and developed. A simple algorithm of the crystal loop detection and preliminary recognition has been designed and developed. The crystal loop detection algorithm is based on finding out the crystal loop ending point (opposite to the crystal loop pin) using image cross section (digital image column) profile analysis. The crystal loop preliminary recognition procedure is based on finding out the crystal loop sizes and position using image cross section profile analysis. The crystal loop fine recognition procedure based on Hooke-Jeeves pattern search method with an ellipse as a fitting pattern has been designed and developed. The procedure of restoring missing coordinate of the crystal loop is described. Based on developed algorithms and procedures the optimal auto-centering procedure has been designed and developed. A procedure of optimal manual crystal centering (Two Clicks Procedure) has been designed and developed. Developed procedures have been integrated into control software system PCCS installed at crystallography beamlines Photon Factory BL5A and PF-AR NW12, KEK.

Keywords: X-ray crystallography; high-throughput protein crystallography; control system; SR; object detection and recognition; Linux

* Corresponding author. Tel.: +7-913-065-8195.
  E-mail address: yurii3000@gmail.com

1. Introduction

Crystal centering procedure is important part of X-ray protein crystallography experiment. A typical protein crystal holder is tiny nylon loop with a size about 50÷500μm. A typical protein crystal size is about 40÷400μm. It takes several minutes to center manually protein crystal under investigation after mounting crystal holder onto X-ray beam area. In contrast, to center the crystal automatically during the X-ray high-throughput protein crystallography experiment becomes not simple procedure. There are several steps of centering the crystal as it was described by Karain et al. (2002) and Pauluhn et al. (2011). Steps typically include several consequent detections, recognitions, rotations around goniometer axis and shifts along coordinate axes using data analysis of the X-ray beam area optical images with different magnification of monitoring optical system. In the case, when the crystal size is about the crystal loop size, centering procedure is completed after centering the crystal loop. In the case, the crystal is smaller than the crystal loop, crystal centering procedure is started, which is similar to crystal loop centering procedure but more complicated and less reliable.

The step of detection and recognition of the crystal loop and the crystal in the loop is crucial in the centering procedure. There are several approaches to solve the problem. One of it, as it was described by Roth et al. (2002) and Jain et al. (2007), is image analysis using digital filters to reduce the image noise and to determine image pixel's intensity gradient or standard deviation of the intensity to detect crystal loop, crystal edges and crystal corners. Another approach, as it was described by Pauluhn et al. (2011), is calculation of ratio of black and white pixels of preliminary converted image onto black and white mode to find the boundaries of the crystal loop. Nevertheless all algorithms have some percentage of fault cases of centering and the problem of crystal loop and crystal centering still is actual.

In this report we describe simple algorithm of crystal loop detection and preliminary recognition based on profile analysis of optical image cross-section curves. A procedure of crystal loop fine recognition using Hooke-Jeeves pattern search method with an ellipse as a fitting pattern has been designed and developed. An optimal manual crystal centering procedure (Two Clicks Procedure) is described.

2. Crystal loop centering

Schematically the optical image of the X-ray beam area is defined as shown at Fig.1a. The goniometer is on the left side of the optical image of X-ray beam area. Goniometer axis G_{axis} coincides with axis X of optical image of the X-ray beam area and crosses X-ray beam. Coordinate axes X and Y of the X-ray beam area coincides with coordinate axes X and Y of the optical image. Coordinate origin of the X-ray beam area coincides with coordinate origin of the optical image and is in the center of the optical image. The X-ray beam crosses coordinate origin of the optical image.

Crystal loop centering algorithm is based on finding coordinates (x_{cp}, y_{cp}, z_{cp}) of some centering point of the crystal loop. Two coordinates (x_{cp}, y_{cp}) can be obtained using the optical image of the X-ray beam area directly by clicking the mouse pointer onto the centering point of the crystal loop (manual centering mode) or automatically (auto-centering mode). Unknown centering point coordinate z_c can be obtained after solving the geometric task with a set of two XY point coordinates before and after rotation the crystal loop by angle α=90° around the axis X using right screw along the positive direction of the axis X, as it is described in Korn and Korn (1968):

$$\begin{pmatrix} x_{cpr} \\ y_{cpr} \\ z_{cpr} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_{cp} \\ y_{cp} \\ z_{cp} \end{pmatrix} = \begin{pmatrix} x_{cp} \\ -z_{cp} \\ y_{cp} \end{pmatrix}, \tag{1}$$

where (x_{cp}, y_{cp}, z_{cp}), $(x_{cpr}, y_{cpr}, z_{cpr})$ – centering point coordinates before and after the crystal loop rotation. Other two coordinates (x_{cpr}, y_{cpr}) are obtained in the same manner as coordinates (x_{cp}, y_{cp}) at previous step. The unknown coordinate z_c is defined from (2):

$$z_{cp} = -y_{cpr}, \tag{2}$$