



FOUNDATIONS
ADVANCES

Volume 77 (2021)

Supporting information for article:

**A reference-based multi-lattice indexing method integrating
prior information correction and iterative refinement in protein
crystallography**

**Qiang Zhou, Zeng-Qiang Gao, Zheng Dong, Yu-Meng Jiang, Zhun She,
Zhi Geng and Yu-Hui Dong**



Figure S1 A view of four lysozyme crystals placed in a cryo-loop, taken by an on-axis camera.

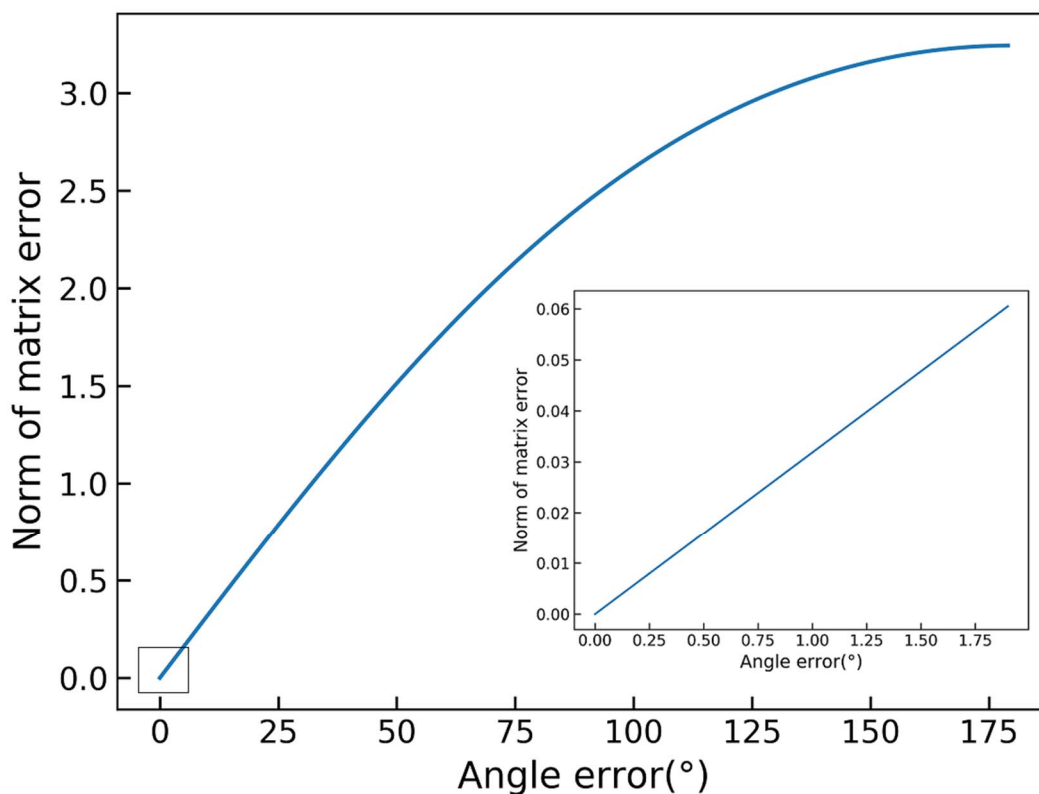


Figure S2 Relationship of the orientation matrix norm error against error of single-axis rotation angle. Inset plots an enlarged view with small matrix norm error, scaling almost linearly with single-axis rotation error.

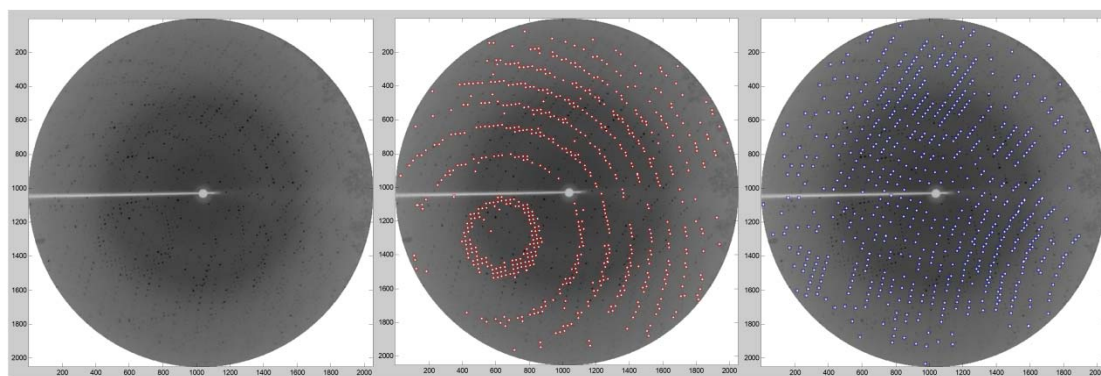


Figure S3 Indexing result of the two-crystal experimental data. Two identified lattices are shown on the original image with predicted boxes overlaid.

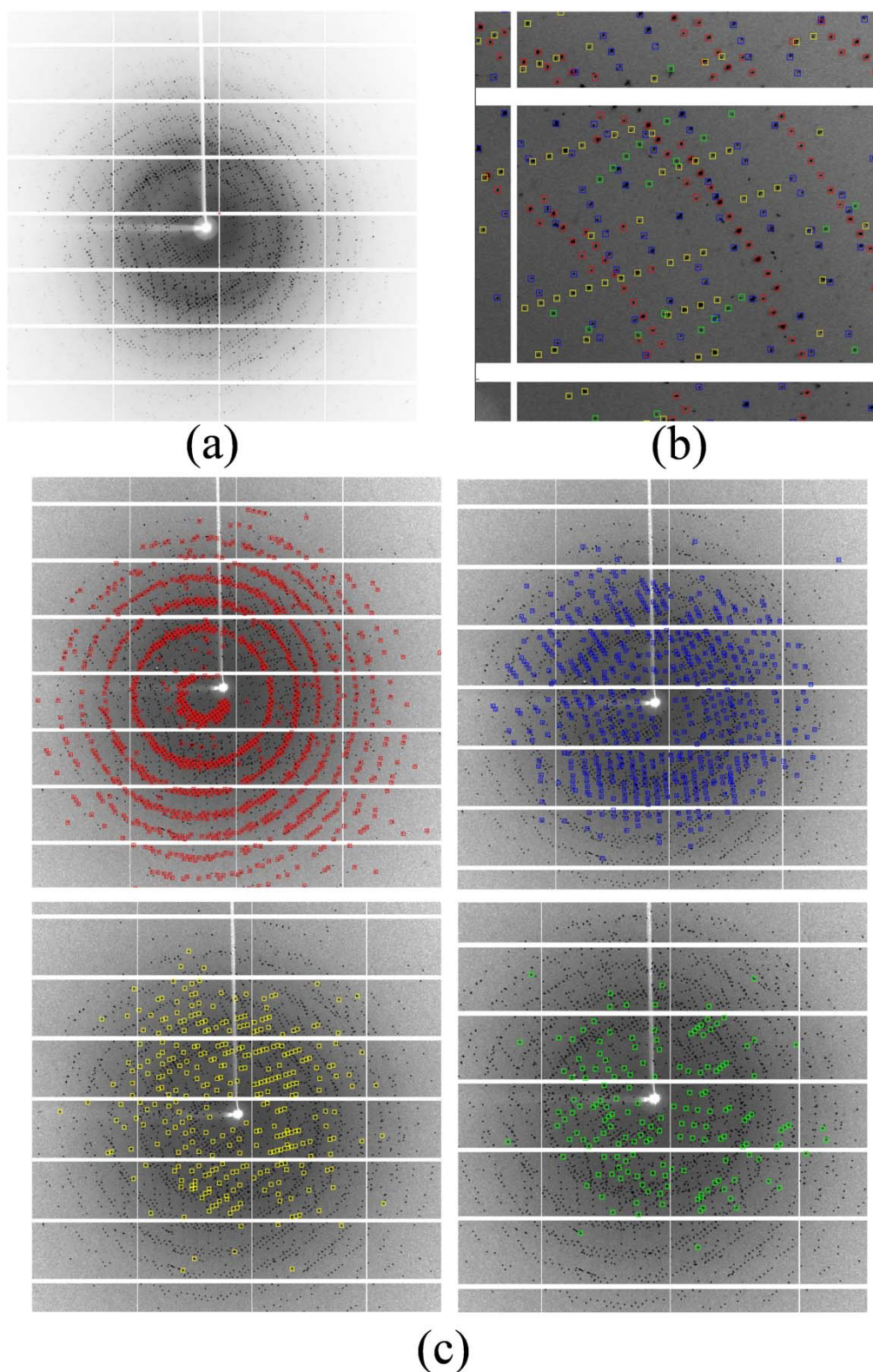


Figure S4 Experimental test result using four-crystal diffraction data. (a) Original measured diffraction image used for indexing. (b) Enlarged view of a simultaneous display of four identified lattices with differently colored prediction boxes overlaid. (c) Separate ensembles of observed spots belonging to each crystal lattice boxed with predicted positions.

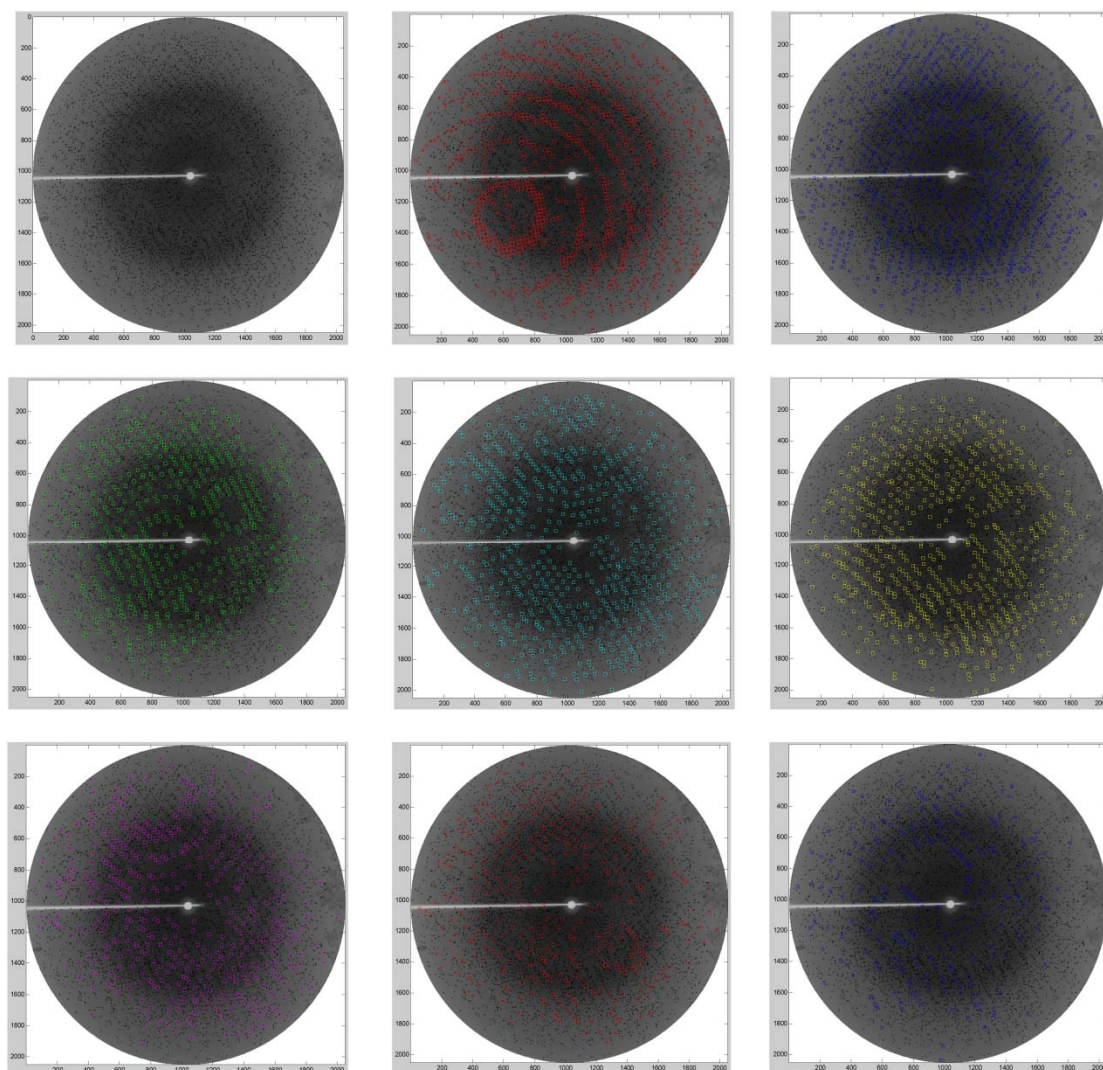


Figure S5 Indexing result using semi-synthetic eight-lattice diffraction image. Shown here are the original synthetic image and eight distinguished lattices.

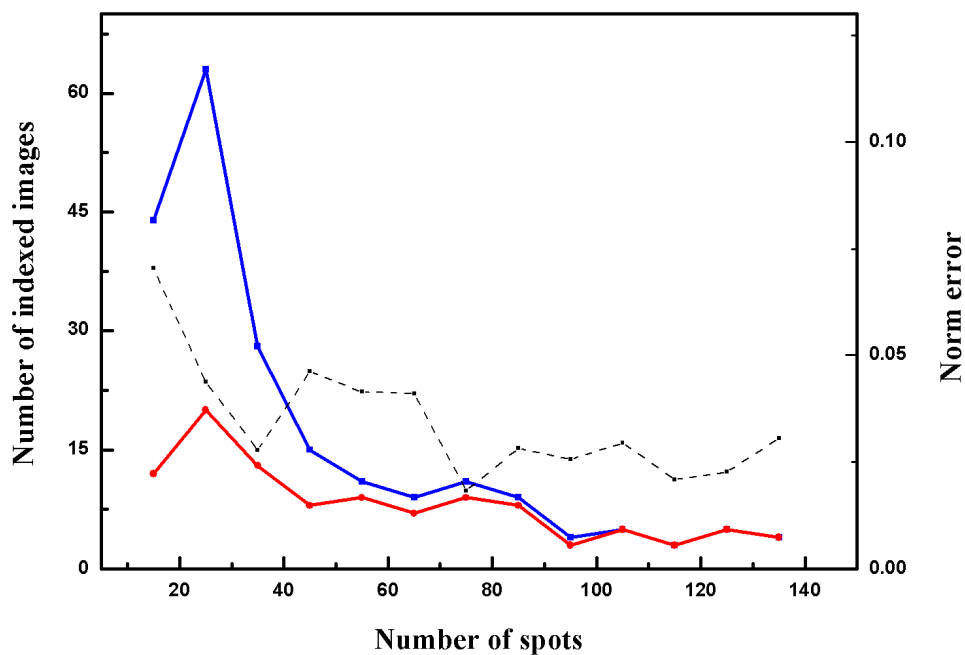


Figure S6 Test on SFX lysozyme data set containing 500 diffraction images. Shown here is the number of successfully indexed images using our multi-lattice indexing method (Blue line) and *MOSFLM* (Red line) as a function of the number of binned observed spots in each image (note that the binning interval is ten spots), respectively. Additionally, the gray dotted line reveals the average norm error of orientation matrices derived from the commonly indexed images between *MOSFLM* and the proposed method.

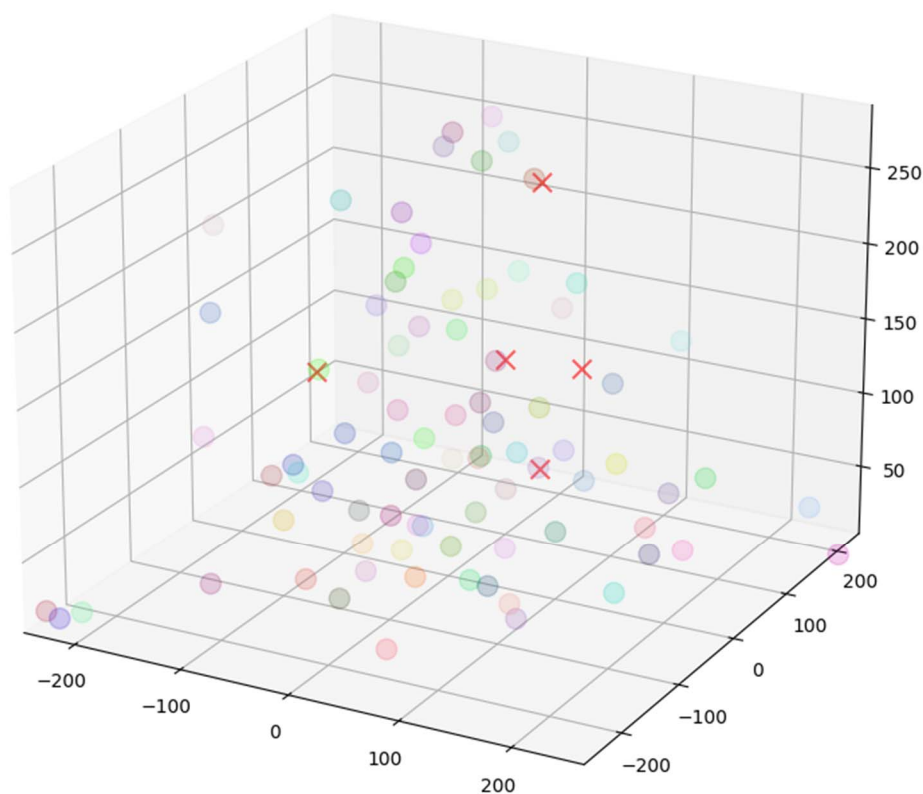


Figure S7 Hierarchy clustering result of all candidate orientations with match-rate higher than 0.15 after the first iteration of multi-lattice indexing (without a further retry-after-subtract process) using a synthetic diffraction image containing five crystal lattices (total 82 clusters). Cross symbols indicate the five known real orientations. It should be noted that the 3D space coordinate of each orientation cluster shown in this image is represented as $(n_x \ n_y \ n_z) * \theta$.

Table S1 Execution times of indexing all experimental data sets in this work on a personal computer equipped with Intel Core Processor i7-9700K @ 3.6GHz.

Experimental data set	No. of images	Resolution cutoff (Å)	Total execution time (seconds)
SR two-lattice data set	1	5.0	23
SR three-lattice data set	1	5.0	37
SR four-lattice data set	1	5.0	74
SR eight-lattice data set	1	5.0	197
SR twinned data set	1	5.0	73
SFX data set	500	3.0	1509

Table S2 A comparison result of indexing SFX lysozyme data set including 500 images with *MOSFLM* and the proposed multi-lattice indexing method.

Method	Total images	Indexed images	Indexed lattices	Images of one lattice	Two lattices	Three lattices
<i>MOSFLM</i>	500	152	152	152	0	0
This work	500	260	313	216	35	9