Electronic Supplementary Material (ESI) for CrystEngComm.
This journal is © The Royal Society of Chemistry 2017

Supplementary Material

Combine Cross-Diffusion Microbatch Method with Seeding Technique to Enhance Protein Crystallization Based on a Common Dispersing Agent

Hai Hou [1‡] Bo Wang [1a‡] Peng-Peng Xie [1] Yun-Zhu Guo[2] Jin Li[1] and Da-Chuan Yin [1*]

[1]*Key Laboratory for Space Bioscience & Biotechnology, School of Life Sciences, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, P. R. China and*

[2] *Chang'an University, Middle-section of Nan'er Huan Road, Xi'an 710064, Shaanxi, P. R. China*

[a]*Present Address: School of Life Sciences, Peking University, Beijing 100871, P. R. China*

**Correspondence e-mail: yindc@nwpu.edu.cn*

[‡] *These authors contributed equally to this work.*

Supplementary Table S1. The condition of different protein in crystallization trials.

Protein	buffer	Initial concentrations
lysozyme	0.1 M sodium acetate pH 4.60	20 mg ml^{-1}
proteinase K	0.05 M sodium cacodylate, 0.08 M magnesium acetate pH 6.50	20 mg ml^{-1}
concanavalin A VI	0.1 M bis-tris pH 6.5	8 mg ml^{-1}
catalase	0.1 M succinic acid pH 7.00	8 mg ml^{-1}
thaumatin	0.1 M potassium phosphate pH 7.00	20 mg ml^{-1}
HTA	20 mM Tris-HCl pH 8.0	2 mg ml^{-1}
glucose isomerase	25 mM HEPES sodium buffer pH 7.00	8 mg ml^{-1}

In crystallization trials, different proteins were dissolved in different buffers as list above. These buffers were chosen as the best suit buffer according to different protein. After dissolved, the protein solutions were mixed with the screening kit Index™ at a ratio of 1 μl: 1 μl by crystallization robot (Screenmaker; Innovadyne Technologies Inc., USA). The CDM crystallization plate was then sealed with Crystal Clear Tape (Hampton Research, catalogue No. HR4-506) and placed in a temperature controller at 293 K.

Supplementary Table S2. Number of hit conditions with no seeds added and with seeds (concentrated seed stocks) using the CDM crystallization plate.

Proteins	No. of hits from 96 conditions without seeding	No. of hits from 96 conditions in first seeding	No. of hits from 96 conditions in second seeding
Lysozyme	21	41	49
Proteinase K	50	76	82
Concanavalin A VI	45	49	48
Catalase	24	27	28
Glucose isomerase	36	56	51
Thaumatin	2	4	4
HTA	28	33	35

Each protein was initially screened in the CDM crystallization plate using Index™ kits, and the single crystal with the best morphology was selected for making the seed stock. We tested the seed screening method with the seed stock dilutions in a pre-equilibrated CDM crystallization plate and compared it with the initial crystallization screening. After the crystals had grown in two crystallization plates, the images from all droplets on the plates were examined using UV light to exclude salt crystals. Then, the single crystal with the best morphology from the seeding plate was selected for a second seeding trial to validate the crystallization reproducibility. The total number of hits with and without seeding for each tested protein is summarized in Table S2. Compared with the initial crystallization screening without seeding, the seeding method in the CDM crystallization plate clearly improved the crystallization screening hits for the 7 tested protein crystals.

Supplementary Table S3. X-ray diffraction data statistics of five proteinase K crystals

c	Diffraction data statistics of proteinase K crystals with no seeding				
	Crystal 1	Crystal 2	Crystal 3	Crystal 4	Crystal 5
Resolution range (Å)	50-2.01 (2.06-2.01)	50-2.06 (2.12-2.06)	50-2.13 (2.19-2.13)	50-2.17 (2.24-2.17)	50-2.18 (2.25-2.18)
Mosaicity (°)	0.73	0.76	0.80	0.98	0.80
$\langle I \rangle / \langle \sigma(I) \rangle$	22.05 (2.07)	12.79(2.08)	10.43 (2.03)	8.40 (2.65)	7.26 (2.30)
Completeness (%)	98.8(93.5)	97.5(92.1)	98.8 (90.2)	99.8(94.6)	96.9(93.2)
R_{merge} [a] (%)	11.40(41.2)	12.2(54.3)	12.4 (45.1)	13.00(36.8)	12.30(52.7)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	$a = 67.83$ $b = 67.83$ $c = 102.28$	$a = 67.75$ $b = 67.75$ $c = 101.12$	$a = 67.72$ $b = 67.72$ $c = 102.23$	$a = 67.84$ $b = 67.84$ $c = 101.79$	$a = 67.87$ $b = 67.87$ $c = 102.21$
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5				
Total frames	90				
Oscillation angle (°)	1				
Detector distance (mm)	150				
Wavelength (Å)	1.54179				

Values in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Table S4. X-ray diffraction data statistics of five proteinase K crystals

	Diffraction data statistics of proteinase K crystals with the first seeding				
	Crystal 1	Crystal 2	Crystal 3	Crystal 4	Crystal 5
Resolution range (Å)	50-1.82 (1.86-1.82)	50-1.86 (1.90-1.86)	50-1.87 (1.91-1.87)	50-1.90 (1.94-1.90)	50-1.91 (1.95-1.91)
Mosaicity (°)	0.56	0.64	0.51	0.76	0.72
$\langle I \rangle / \langle \sigma(I) \rangle$	21.78 (2.14)	21.09 (2.07)	10.72 (2.16)	22.36 (2.02)	15.21 (2.08)
Completeness (%)	99.4(95.1)	99.2(92.4)	97.4(93.8)	97.6(91.5)	97.1 (95.3)
R_{merge} [a] (%)	8.50(34.6)	8.40(46.5)	8.00(41.6)	7.30(41.2)	10.7 (45.9)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	*a* = 67.59 *b* = 67.59 *c* = 101.29	*a* = 67.24 *b* = 67.24 *c* = 100.78	*a* = 67.76 *b* = 67.76 *c* = 101.53	*a* = 67.86 *b* = 67.86 *c* = 101.47	*a* = 67.49 *b* = 67.49 *c* = 101.04
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5				
Total frames	90				
Oscillation angle (°)	1				
Detector distance (mm)	150				
Wavelength (Å)	1.54179				

Values in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Table S5. X-ray diffraction data statistics of five proteinase K crystals

	Diffraction data statistics of proteinase K crystals with second seeding				
	Crystal 1	Crystal 2	Crystal 3	Crystal 4	Crystal 5
Resolution range (Å)	50-1.78 (1.81-1.78)	50-1.78 (1.81-1.78)	50-1.79 (1.82-1.79)	50-1.82 (1.86-1.82)	50-1.83 (1.87-1.83)
Mosaicity (°)	0.78	0.57	0.54	0.58	0.50
$<I>/<\sigma(I)>$	24.80 (2.10)	14.93 (2.09)	32.28 (2.01)	13.29 (2.26)	11.54 (2.23)
Completeness (%)	99.1(95.2)	98.9(95.2)	99.2 (88.7)	99.4(93.4)	98.7(94.2)
R_{merge} [a] (%)	8.70(48.2)	8.47(51.2)	7.6 (48.0)	12.90(34.3)	7.30(42.4)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	*a* = 67.58 *b* = 67.58 *c* = 101.26	*a* = 67.54 *b* = 67.54 *c* = 101.72	*a* = 76.47 *b* = 76.47 *c* = 101.04	*a* = 67.41 *b* = 67.41 *c* = 100.98	*a* = 67.80 *b* = 67.80 *c* = 102.44
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5				
Total frames	90				
Oscillation angle (°)	1				
Detector distance (mm)	150				
Wavelength (Å)	1.54179				

Values in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Table S6. X-ray diffraction data statistics of five lysozyme crystals

	Diffraction data statistics of Lysozyme crystals with no seeding				
	Crystal 1	Crystal 2	Crystal 3	Crystal 4	Crystal 5
Resolution range (Å)	50-1.90 (1.94-1.90)	50-1.90 (1.94-1.90)	50-1.92 (1.96-1.92)	50-1.93 (1.97-1.93)	50-1.93 (1.97-1.93)
Mosaicity (°)	1.15	1.27	1.27	0.99	0.66
$<I>/<\sigma(I)>$	17.13 (2.03)	27.85 (2.35)	16.98 (2.06)	14.63 (2.12)	21.55 (2.01)
Completeness (%)	99.6(95.4)	99.3 (95.3)	98.7(93.4)	99.3(94.6)	99.5(93.8)
R_{merge} [a] (%)	10.70(37.9)	9.2 (41.4)	9.30(44.0)	10.40(34.7)	11.10(37.8)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	*a* = 76.64 *b* = 76.64 *c* = 36.94	*a* = 76.28 *b* = 76.28 *c* = 36.75	*a* = 76.65 *b* = 76.65 *c* = 37.42	*a* = 76.38 *b* = 76.38 *c* = 38.20	*a* = 76.15 *b* = 76.15 *c* = 37.00
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5				
Total frames	90				
Oscillation angle (°)	1				
Detector distance (mm)	150				
Wavelength (Å)	1.54179				

Values in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Table S7. X-ray diffraction data statistics of five lysozyme crystals

	Diffraction data statistics of Lysozyme crystals with the first seeding				
	Crystal 1	Crystal 2	Crystal 3	Crystal 4	Crystal 5
Resolution range (Å)	50-1.73 (1.76-1.73)	50-1.76 (1.79-1.76)	50-1.77 (1.80-1.77)	50-1.85 (1.89-1.85)	50-1.93 (1.97-1.93)
Mosaicity (°)	0.53	0.46	0.85	0.57	0.65
$\langle I \rangle / \langle \sigma(I) \rangle$	12.06 (2.09)	11.79 (2.04)	23.49 (2.05)	14.93 (2.09)	25.92 (2.02)
Completeness (%)	96.8(94.2)	99.6(96.4)	95.9 (87.3)	97.4(92.6)	99.8(95.6)
R_{merge} [a] (%)	6.70(35.6)	7.30(30.6)	7.10 (32.9)	10.30(47.4)	8.30(35.8)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	$a = 78.02$ $b = 78.02$ $c = 37.55$	$a = 76.47$ $b = 76.47$ $c = 37.58$	$a = 78.12$ $b = 78.12$ $c = 36.99$	$a = 76.71$ $b = 76.71$ $c = 37.32$	$a = 76.72$ $b = 76.72$ $c = 37.51$
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5				
Total frames	90				
Oscillation angle (°)	1				
Detector distance (mm)	150				
Wavelength (Å)	1.54179				

Values in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Table S8. X-ray diffraction data statistics of five lysozyme crystals

	Diffraction data statistics of Lysozyme crystals with the second seeding				
	Crystal 1	Crystal 2	Crystal 3	Crystal 4	Crystal 5
Resolution range (Å)	50-1.74 (1.77-1.74)	50-1.75 (1.78-1.75)	50-1.78 (1.81-1.78)	50-1.79 (1.82-1.79)	50-1.88 (1.92-1.88)
Mosaicity (°)	0.44	0.76	0.45	0.54	0.78
$\langle I \rangle / \langle \sigma(I) \rangle$	33.65 (2.24)	28.09 (2.11)	18.92 (2.02)	32.28 (2.01)	31.11 (2.05)
Completeness (%)	99.3(94.6)	99.6(96.7)	98.6(95.4)	99.2 (88.7)	98.8(94.7)
R_{merge} [a] (%)	6.80(31.6)	9.40(38.5)	7.70(35.6)	7.6 (48.0)	6.30(29.8)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	*a* = 76.84 *b* = 76.84 *c* = 36.83	*a* = 76.19 *b* = 76.19 *c* = 36.84	*a* = 77.01 *b* = 77.01 *c* = 36.99	*a* = 76.47 *b* = 76.47 *c* = 37.06	*a* = 76.97 *b* = 76.97 *c* = 36.97
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5				
Total frames	90				
Oscillation angle (°)	1				
Detector distance (mm)	150				
Wavelength (Å)	1.54179				

Values in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Table S9. X-ray diffraction data statistics of three thaumatin crystals. We diffracted the single crystals with the best morphology for each group (without seeding, first seeding, second seeding). As a result, in term of resolution limits, mosaicity and $R_{merge,}$ the crystals after seeding demonstrate a better state compare to crystal without seeding.

	Diffraction data statistics of thaumatin crystals in crystallization trails		
	Without Seeding	First Seeding	Second Seeding
Resolution range (Å)	50-2.78 (2.83-2.78)	50-2.28 (2.35-2.28)	50-2.21 (2.28-2.21)
Mosaicity (°)	2.5	1.08	1.26
$<I>/<\sigma(I)>$	12.55 (2.19)	17.6 (2.10)	11.44 (2.14)
Completeness (%)	99.5 (98.8)	97.1 (89.3)	90.9 (94.1)
R_{merge} [a] (%)	19.4 (44.8)	10.6 (34.7)	11.8 (34.2)
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions *a*, *b*, *c* (Å)	*a* = 56.98 *b* = 57.43 *c* = 149.25	*a* = 57.66 *b* = 57.65 *c* = 149.89	*a* = 57.64 *b* = 58.14 *c* = 161.90
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Exposure time (min)	5		
Total frames	120		
Oscillation angle (°)	1		
Rotation range (°)	150		
Detector distance (mm)	200		
Wavelength (Å)	1.54179		

Values in parentheses are for the highest resolution shell.

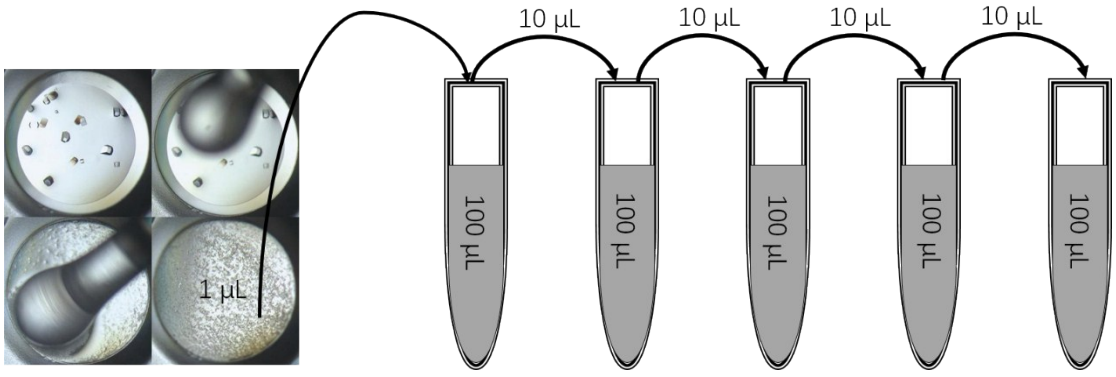
[a] $R_{merge} = \sum_h \sum_i |I_{(h)} - I_{i(h)}| / \sum_h \sum_i I_{i(h)}$, where $I_{i(h)}$ and $I_{(h)}$ are the *i*th and mean measurement of reflection *h*.

Supplementary Figure S1. Each protein was initially screened in the CDM crystallization plate using Index™ kits, and the single crystal with the best morphology was selected for making the seed stock. After obtain all kinds of protein crystals, we tested stability of crystal.

At first, we mixed the components of candidate of "common dispersing agent" in a little serum bottle (1ml), and then we put the crystal which we pick it up from the optimal condition to the little serum bottle containing the common dispersing agent. Then, we can observe if morphological changes occurred on the crystals and take photo using camera connected to microscope. At start, we took one photo per 10 seconds, and after 5 minutes, we took one photo per minute, after 30 minutes, we took one photo per 10 minutes, after 3 hours, took one photo per hour, after 10 hours, took one photo per day until 3 days. Eventually, we determined candidate of the common dispersing agent through screening several single agents and mixtures.

Next, The protein solutions were mixed with the screening kit Index™ at a ratio of 1 µl: 1 µl by machine and added in the CDM crystallization plate to construct new pre-equilibrated CDM plate.

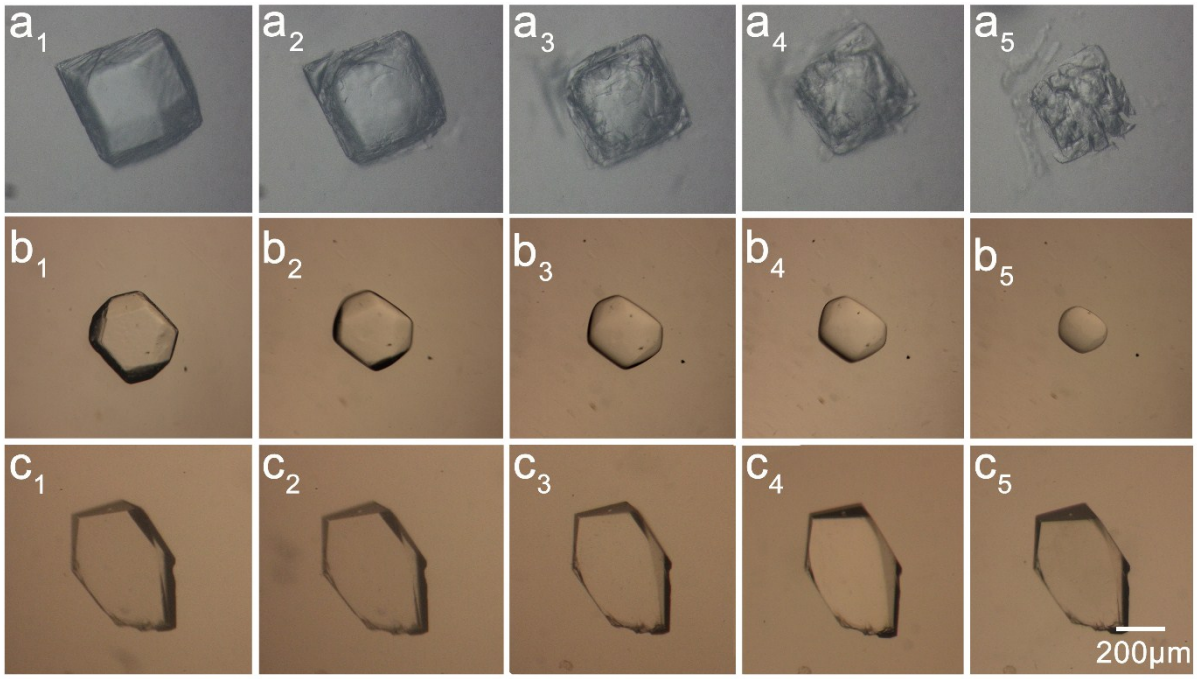
After that, we crushed the crystals obtained at start and diluted at different folds according to the protocol as below.



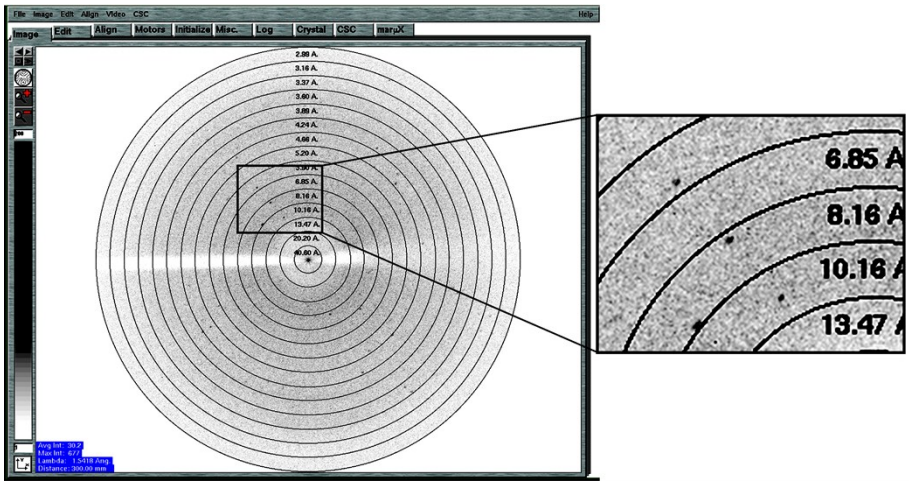
We crushed obtained crystals into micro-crystals with 1 µL common dispersing agent by Crystal crusher (Hampton Research Inc.) as above. Next, we aspirated this 1 µL droplet containing crushed crystals and added into 100 common dispersing agent, we called it

“Sample 1”. Furthermore, we aspirated 10 μL from Sample 1 to add into another 100 μL common dispersing agent and called “Sample 2”. According to this method, we made a series of concentration gradient sample: Sample 1, Sample 2, Sample 3,…. Each different concentration sample will be add to a pre-equilibrium CDM plate with 0.1 μL per droplet. Then the plate was sealed and put into a temperature controller at 273 K.

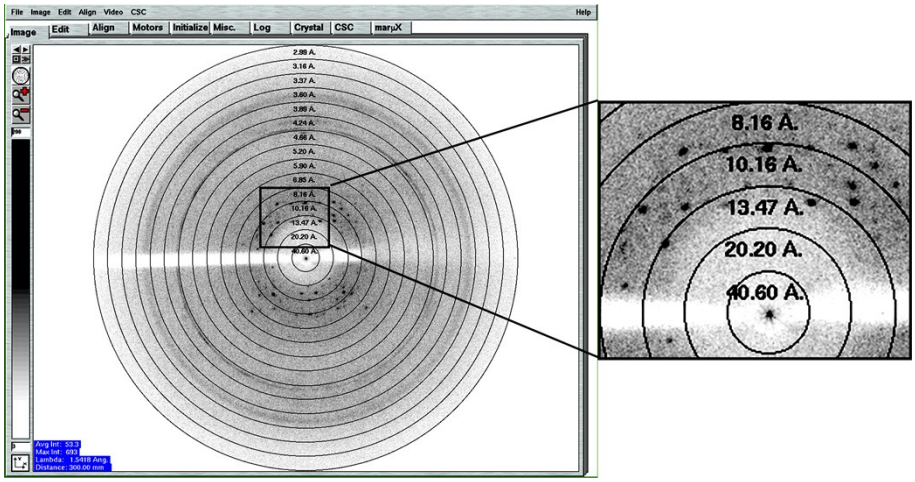
Supplementary Figure S2. Lysozyme protein crystals exhibited different status on morphology in different chemical agents with time. (a_1), (b_1) and (c_1) showed the initial morphology of lysozyme protein crystals in PEG10000, sodium chloride, and universal dispersing agent solutions, respectively. (a_2)-(a_5) represented the crystal changes in the PEG10000 within 1 minute, 3 minutes, 10 minutes, 20 minutes, the crystal dissolved and cracked quickly as time goes on. (b_2)-(b_5) showed the crystal defects in the sodium chloride within 1 minutes, 5 minutes, 10 minutes, 20 minutes, the crystal dissolved quickly. (c_2)-(c_5) exhibited the morphology in the universal dispersing agent we used within 1 minute, 5 minutes, 60 minutes, 300 minutes. These results showed that the universal dispersing agent has insolubility for lysozyme protein crystals.



Supplementary Figure S3. The diffraction image of concanavalin A VI protein crystal after seeding. Crystals with the best morphology were selected before and after seeding, and the crystals were examined in the X-ray single crystal diffraction system. Crystal after seeding diffracts some spots in diffraction pattern. Compared to concanavalin A VI crystal after seeding, the crystal before seeding cannot diffract any spots (data not shown). Therefore, concanavalin A VI crystal after seeding demonstrated improved quality than before seeding.



Supplementary Figure S4. The diffraction image of HTA protein crystal after seeding. Similar to the previous operation, HTA protein crystals with the best morphology were selected before and after seeding were examined in the X-ray single crystal diffraction system. The result is that crystal before seeding cannot diffract any spots (data not shown), on the contrary, the crystal after seeding diffracted lots of clear spots. As a result, HTA crystal after seeding exhibited better diffraction quality than before seeding.



Supplementary Figure S5. The diffraction image of glucose isomerase protein crystal before and after seeding. Similar to the previous operation, we selected the crystals with the best morphology, and the crystals were examined in the X-ray single crystal diffraction system. The crystal without seeding just got a few diffraction spots, however, the crystal with seeding diffracted more diffraction spots dramatically (nearly reach to 3 Å in resolution). Compared to glucose isomerase protein crystal before seeding, the crystal after seeding clearly demonstrated a dramatic quality improvement on diffraction. (a) Diffraction image of the glucose isomerase protein crystals before seeding. (b) Diffraction image of the glucose isomerase protein crystals after seeding.