

Volume 52 (2019)

Supporting information for article:

*Aquarium*: an automatic data-processing and experiment information management system for biological macromolecular crystallography beamlines

Feng Yu, Qisheng Wang, Minjun Li, Huan Zhou, Ke Liu, Kunhao Zhang, Zhijun Wang, Qin Xu, Chuanyan Xu, Qiangyan Pan and Jianhua He

Filter by data name (case insenaltive)					Auto Refresh	ł.											Select date range		Open BL17U1			
Data name il	Spacegroup If	+11	6 II	-11	- 11	p.11	y II	Oscwidth If	Frames If	Resolution []	Inner Rmeas If	Outer Rmeas II	Rmeas II	Inner AnomCC II	AnomCC II	Anom Limit Res II	Lotat II	Score II	Residues(built) If	Built Ratio II	NCS IT	Time II
puck9_16	P 21 21 2	121-30	214.07	33.93	90.00	90.00	90.00	0.20	1800	2.10	0.092	1.655	0.205	+0.101	0.236		0.304	199.859				2018-05-03 17:07 GMT+6
suck9_15	# 21 21 2	121.10	213.64	33.85	90.00	90.00	90.00	1.00	360	1.65	0.046	1.243	0.078	0.096	0.033		0.495	886.290				2018-05-03 17:01 GMT+8
uck9_14	P 21 21 2	121.72	213.81	33.65	90.00	90.00	90.00	1.00	360	1.64	0.047	2.129	0.093	+0.039	0.049		0.506	593.683				2018-05-03 16:54 GMT+8
uck9_13	P 21 21 2	122.04	213.78	23.97	90.00	90.00	90.00	1.00	360	1.70	0.053	1.416	0.093	-0.003	+0.051		0.505	647.914				2018-05-03 16:47 GMT+8
web9_12	P 21 21 2	122.01	214.23	33.96	90.00	90.00	90.00	1.00	360	1.75	0.071	1.617	0.102	0.662	0.362		0.501	601-245				2018-05-03 15:40 GMT+8
uck13_5	P 21 21 21	63.62	65.01	110.73	90.00	90.00	90.00	1.00	360	1.94	0.062	1.543	0.158	0.072	0.077		0.456	347.179				2010-05-03 16:33 GMT+0
uck13_3	# 21 21 21	64.18	85.50	110.75	90.00	90.00	90.00	1.00	360	1.55	0.039	1.524	0.106	0.083	0.001		0.485	825.993				2018-05-03 16:24 GM7+8
uck12_16	P 2 2 21	85.65	108.22	64.33	90.00	90.00	90.00	1.00	360	1.74	0.149	0.905	0.334	-0.008	0.615		0.481	244.002				2018-05-03 16:17 GMT+8
zek12_15	P 21 21 2	85.64	108.74	64.27	90.00	90.00	90.00	1.00	360	1.64	0.124	1.617	0.262	+0.037	0.062		0.476	314.364				2018-05-03 16:10 GMT+8
eck12_14	P 21 21 2	108.35	64.42	85.27	90.00	95.00	90.00	1.60	360	1.57	0.083	1.932	0.156	0.552	0.112		0.435	575.410				2018-05-03 16+03 GMT+8
vck12_12	P 21 21 21	64.11	65.65	110.06	90.00	90.00	90.00	1.00	360	1.65	0.055	1.616	0.125	0.296	-0.001		0.456	595.125				2018-05-03 15:57 GMT+8
ack12_12	# 2 2 2	64.28	85.73	108,45	90.00	90.00	90.00	1.00	360	3,90	0,102	1.271	0.24	0.039	-0,124		Q.458	229.342				2018-03-03 15:50 GMT+8
uck12_11	P 2 2 21	85.29	110.02	64.21	90.00	90.00	90.00	1.00	360	1.51	0.144	1.696	0.257	-0.001	-0.071		0.457	446.642				2018-05-03 15:41 GMT+8
udk12_10	P 2 2 21	110.40	63.90	85.68	90.00	90.00	90.00	1.00	360	1.87	0.096	1.524	0.208	0.032	-0.014		0.503	264.657				2018-05-03 15:35 GMT+8
uck12_9	P 21 21 21	64.46	85.44	111.05	90.00	90.00	90.05	1.00	360	1.47	0.042	1.385	0.103	0.029	-0.024		0.501	1,008.016				2018-05-03 15-28 GMT+8
rsk12_0	P 21 21 2	05.67	111.05	64.05	90.00	90.00	90.00	1.00	360	1.30	0.042	0.962	0.078	-0.17	-0.053		0.485	1.326.530				2010-05-03 15:21 GMT+0
ush12_6	P 21 21 21	64.50	85.91	111.46	90.00	90.00	90.00	1.00	360	2.47	0.048	1.292	0.104	-0.091	0.019		0.493	1.027.355				2018-05-03 15:11 GMT+8
ack13_2	P 21 21 2	06-30	111-12	64.46	90.00	90.00	90.00	1.00	360	2.00	0.066	1.569	0.177	0.265	0.046		0.475	270.545				2018-05-03 14:49 GMT+8
uck13_1	P 2 2 2	64.02	85.88	111.40	90.00	90.00	90.00	1.00	360	1.96	0.076	1,658	0.273	+0.353	0.035		0.456	237.608				2018-05-03 14:39 GM7+8
uck13_16	P 21 21 2	247.21	80.27	127.06	90.00	90.00	90.00	1.00	360	5.99	0.065	1.804	0.11	0.549	0.3		0.509	14.309				2018-05-03 14:25 GMT+8
ck13_13_set2								1.00	360	7.34												2018-05-03 14:02 GMT+0
2194_52_52.6E								0.20	3600	7,40												2018-05-03 13:53 GMT+8
uck13_13	9222	80.15	127.14	247.26	90.00	90.00	90.00	0.20	1800	6.20	0.097	1.792	0.21	+0.372	-0.156		0.462	3.941				2018-05-03 13:45 GMT+8
uck13_9	P 2 2 21	127.58	245.88	80.37	90.00	90.00	90.00	0.50	720	6.39	0.086	1.297	0.183	+0.308	0.013		0.507	9.490				2018-05-03 13:19 GMT+8
puck9_11	P 21 21 2	121.60	213.77	33.88	90.00	90.00	90.00	0.50	720	1.76	0.03	1.782	0.063	+0.095	-0.033		0.503	805.614				2018-05-03 12:

First Previous 1 2 Next Last

Figure S1. A summary of datasets of a visit.  $R_{meas}$ , AnormCC and Anorm Limit Res come from Aimless; Lstat is from CTRUNCATE; scores were calculated using the number of unique reflections, the completeness, the number of molecules per asymmetric unit,  $I/\sigma$  and the volume of unit cell. The scoring function is identical with that adopted in XChemExplorer; the built residues, built ratio and NCS come from porpoise\_sad.

by data name (case	e insensitive)												Select date n	ange Open
/users/bl17u1/201	80927/ghj/A7/A7	_1_maste	r.h5										Collec	tion time: 2018-09-27 20:0
Sample: A7 Barcode: NA Omega start: 0.000* One wetht: 0.2001			Exposure time: 0.200 s Wavelength: 0.979176 A Distance: 180.00 mm Resolution: 1.36 A 1.13 A									Resolution and spots against rotation		
								-	and a		450		1.7	
									1		350		221 5	
							i i i i	and the second second	Sec. 1		5 300 250		22 000	
nega range: 200	000*		Elux: M		0.074			-	and the second		200		25	
ega range. 360.	.000		Flux, N	• 			-				1	0 50 100 150 200 250 Rotation (degree)	300 350	
of images: 1800	0 images		Beamsi	ze: NA			10000		50	And Address in such as the		Total spots Good spots Re	solution	
ta reduction										Porpois	e(XDS)	rpoise(DIALS) to xia2(3dii)	tite xia2(dials)	utoPROC(XDS)
ta reduction Porpoise(XDS) Space group P 61	Porpoise(DIA a 65.67	■LS) b 65.67	c 131.50	xia2(dials) a 90.00	autoPF β 90.00	ROC(XDS) Y 120.00		Rmerge,	<io> against Batches</io>	Porpois	e(XDS) <b>to solve agenere</b> Solve agenere	rpoise(DIALS)	Completeness and Anon-Completeness	utoPROC(XDS)
ta reduction Porpoise(XDS) Space group P 61	Porpoise(DIA a 65.67	NLS) b 65.67	iia2(3dii) c 131.50	xia2(dials) a 90.00	autoPF β 90.00	ROC(XDS) ¥ 120.00	0.4	Rmerge,	<ir></ir>	7.00 Rnesp	e(XDS)	nation way	Completeness and AconCompleteness a	ganst Resolution
ta reduction torpoise(XDS) Space group > 61	Porpoise(DIA a 65.67 Beam_X	NLS) 5 65.67	c 131.50 Beam_Y	xia2(diats) α 90.00	autoPF β 90.00 Distance	ROC(XDS) Y 120.00	0.4 0.3 0.3 8, 0.2	Rmerge.	≪iα- against Batches	Porpois	e(XDS) the Por	dalin may	Completeness and AcontCompleteness a	utoPROC(XDS) set a final second secon
ta reduction Porpoise(XDS) Space group P 61	Porpoise(DIA a 65.67 Beam_X 155.63	NLS) D 65.67	c 131.50 Beam_Y 163.91	xia2(dials) a 90.00	autoPF β 90.00 Distance 180.00	ROC(XDS) Y 120.00	0.4 0.3 0.3 8 0.2 8 0.2 0.1	Rmerge	<tors< td=""><td>Porpois</td><td>e(XDS) 10110 Por</td><td>dalin way</td><td>Completeness and AcontCompleteness a Completeness and AcontCompleteness a Completeness and AcontCompleteness a Completeness and AcontCompleteness a Completeness and AcontCompleteness a</td><td>utoPROC(XDS) <b>utrative</b></td></tors<>	Porpois	e(XDS) 10110 Por	dalin way	Completeness and AcontCompleteness a Completeness and AcontCompleteness a Completeness and AcontCompleteness a Completeness and AcontCompleteness a Completeness and AcontCompleteness a	utoPROC(XDS) <b>utrative</b>
ta reduction Porpoise(XDS) Space group P 61 start refined	Porpoise(DIA a 65.67 Beam_X 155.63 156.50 0.87	<b>b</b> 65.67	iia2(3dii) c 131.50 Beam_Y 163.91 164.17 0.25	xia2(dials) α 90.00	autoPF β 90.00 Distance 180.00 179.77	кос(XDS) <u>Y</u> 120.00	0.4 0.3 0.3 0.3 0.2 0.11 0.11 0.11	Rinerge	ditr- spinst Batches	Porpois	e(XDS)	talin map	Completeness and AconCompleteness a Completeness a	parter Resolution
ta reduction Porpoise(XDS) Space group P 61 start refined Δ	Porpoise(DIA a 65.67 Beam_X 155.63 156.50 0.87	NLS) : b 65.67	iia2(3dii) c 131.50 Beam_Y 163.91 164.17 0.26	xia2(dials) α 90.00	autoPF β 90.00 Distance 180.00 179.77 -0.23	ROC(XDS) Y 120.00	0.4 0.3 0.2 80 0.2 80 0.2 0.1 0.1 0.0 0.0	Rmerge.	<ir></ir>	Porpois	e(XDS) adjudge de Por	talin map talin map	Completeness and AconCompleteness a Completeness a Completen	utoPROC(XDS) states
ta reduction Porpoise(XDS) Space group P 61 start refined Δ	Porpoise(DIA a 65.67 Beam_X 155.63 156.50 0.87	b 65.67	c 131.50 Beam_Y 163.91 164.17 0.26	xia2(dials) a 90.00	autoPF β 90.00 Distance 180.00 179.77 -0.23	хос(XDS) У 120.00	0.4 0.3 0.2 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1	Rmerge.	<li>dio- against Batches</li> <li>dio- against Batches</li> <li>dio- against Batches</li> <li>800 1000 1200 1400 16</li> <li>Batch</li>	Porpois	e(XDS) adjudie de Por	Additionage	Completeness and AconCompleteness a Completeness a Completen	paint Resultion Competenses Lat 1.69 1.60
a reduction a rorpoise(XDS) space group 2 6 1 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	Porpoise(DIA a 65.67 Beam_X 155.63 156.50 0.87	NLS) : b 65.67	c 131.50 Beam_Y 163.91 164.17 0.26	xia2(dials) a 90.00	autoPF β 90.00 Distance 180.00 179.77 -0.23	ROC(XDS) Y 120.00	0.4 0.2 0.2 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1	Rnsrge,	<li>Alion - against Batches</li> <li>400 - against Batches</li> <li>400 - 1000 1200 1400 16</li> <li>800 1000 1200 1400 16</li> <li>800 1000 1200 1400 16</li>	Porpois	e(XDS) <b>standard</b> Por	tatic maps tatic	Completeness and AconCompleteness a Completeness	and Realize gand Realize and Realize and Realize congeteres concept data tables
a reduction a roturpoise(XDS) Space group > 61 efined A States H	Porpolse(DIA a 65.67 Beam_X 155.63 156.50 0.87 Resolution	NLS) : b 65.67	tia2(3dii) c 131.50 Beam_Y 163.91 164.17 0.26 Observa	a 90.00	autoPF β 90.00 Distance 180.00 179.77 -0.23	ROC(XDS) Y 120.00	04 03 03 03 03 03 03 03 03 03 03 03 03 03	Rmerga,	ion sparst Batches            0           800         1000           Batch         CC(112)	Porpois	Bute spine:	AnomCompleteness	completeness and AnonCompleteness a Completeness a Completeness a Completeness a	and Resolution and Resolution 1.81 1.99 1.60 CC_Anom
tata reduction Porpolse(XOS) Space group P 61 start refined Δ MLESS ell erShell	Porpoise(DIA a 65.67 Beam_X 155.63 156.50 0.87 Resolution 65.75.7.07	b 65.67	tia2(3dii) e 131.50 Beam_Y 163.91 164.17 0.26 Observa 10317	xia2(dials) a 90.00 tions	autoPF           β           90.00           Distance           180.00           179.77           -0.23           U           52	ROC(XDS) Y 120.00	0.3 0.3 0.3 0.2 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	Rmerga,	<li>-(io-spans Bathes</li>	Porpois	e(XDS) a Por	AnomCompleteness 100.00	Completeness and AnonCompleteness and AnonCompleten	aprel Resolution aprel Resolution 1.81 1.49 1.60 CC_Anom 0.12
Ata reduction Porpolse(XDS) Space group P 61 start refined Δ MILESS MILESS Hell nerShell uterShell	Porpoise(DIA a 65.67 Beam_X 155.63 156.50 0.87 Resolution 65.75.7.07 1.62.1.58	b 65.67	ia2(3dii) c 131.50 Beam_Y 163.91 164.17 0.26 Observa 10317 61776	xia2(dials) a 90.00	autoPF           β           90.00           Distance           180.00           179.77           -0.23           U           50           10           51           52           53	ROC(XDS) Y 120.00 nique 26 211	Rmeas 0.060 1.584	Rnerge, 0 200 400 600 1/0(/) 43.50 2.20	-life spars Baldes           -life spars Baldes           -sol two	Porpois	e(XDS) advector Por Both queck	AnomCompleteness 100.00 100.	AnomMultiplicity 10.40 9.60	and Resultion and Resultion Lat 1.49 1.60 CC_Anom 0.12 0.10

Figure S2. The detailed overview of the dataset, including the sample name, wavelength, distance, oscillation width, exposure time, etc. The snapshot, diffraction image and DISTL results show in the right side. Results from data reduction and SAD pipeline can be inspected through dropdown menu. SealWeb applies the thumb icons to represent the quality of data. The more thumbs-up means the better quality of the data (Table S3 of supplementary materials).

## 1. A test example of Porpoise from data reduction to mainchain building

Aquarium runs on a small HPC with Linux system. There are 8 computing nodes, each with two Intel Xeon 2640-V4 CPUs and 64G memory. The storage contains 27 TB flash memory and 48 TB hard disk drive.

A selenium derivatised crystal is used as test example for Porpoise. The dataset contains 1200 frames of raw images. The space group is  $P4_122$ ; the unit cell constants are a = 42.67 Å, b = 42.67 Å, c = 136.02 Å. There are 123 residues in an asymmetric unit, 6 of which are methionine. The data collection, data reduction and SAD phasing token about 390 s, 93.281 s and 104.644 s, respectively. Finally, five Se atoms were found, and 98 of 123 residues were built.

Data collection	1200 frames, omega range $360^{\circ}$ and 0.3
	second/frame.
	360 s for exposure, and extra about 30 s waiting
	for data transport.
Initialization (porpoise_dr)	0.164 s
Spots finding (all frames) and indexing	29.177 s
The first-pass integration as $P1$ (the first 30° of	9.815 s
raw images)	
Laue group determination, the second-pass	44.614 s
integration (all frames), merging and XSCALE	
Systematic absences determination	5.555 s
Reindex and Scaling	2.592 s
Convert intensities to structure factors	1.362 s
Initialization (porpoise_sad)	0.156 s
Checking twinning and anomalous signals	0.785 s
Finding anomalous scatters	44.613 s
Chirality determination	1.939 s
Phasing	2.572 s
Density modification	37.665 s
NCS finding	3.426 s
Model building	13.488s
Total	587.925 s (390s for data collection, 93.281 s for
	data reduction and 104.644 s for SAD phasing and
	model building)

Table S1. The data processing time from data collection to main-chain building

Protein and datas	et information								
The number of resi	idues in the ASU	123							
The number of me	thionine in the ASU	6							
Solvent content		47%							
High-resolution lin	nit	2.64 Å							
The number of uni	que reflections	4155							
High preference of	luster								
Nodes	8 nodes								
Echo node	$2 \times$ Intel Xeon 2640-V4 CPUs (20	$2 \times$ Intel Xeon 2640-V4 CPUs (20 physics cores)							
	64 G memory								
SAD Phasing and	model building								
Programs	Computing resource	Run time	Results						
SHELXC	Run concurrently on a single node	4.150 s	-						
SHELXD	Run concurrently on eight nodes	41.404 s	CFOM = 92.52						
			CCall = 56.19						
			CCweak = 36.33						
			10 Se atoms found.						
SHELXE	Run concurrently on a single node	1.939 s	Contrast:						
			Original hand is 0.412.						
			Inverted hand is 0.367.						
			Pseudo-free CC:						
			Original hand is 53.33 %.						
			Inverted hand is 51.93 %.						
Phaser	Run on a single node	2.572 s	5 Se atoms after refinement						
PARROT and	Run concurrently on eight nodes	37.665 s	PARROT: solvent fraction 60%						
SOLOMON			SOLOMON: solvent fraction						
			50%						
PARROT (NCS)	Run concurrently on a single node	3.426 s	No NCS found						
BUCCANEER	Run concurrently on a single node	13.488 s	98 residues built						

## 2. The icon meaning of data process results in SealWeb

Table S3. The icon meaning of data process results in SealWeb

444	Data reduction: R meas of inner resolution shell $\leq$ 7%, and high-resolution limit $\leq$ 2.0 Å.
	SAD phasing: built residues/estimated residues $\geq 85\%$ .
	Data reduction: R meas of inner resolution shell $\leq 10\%$ , high-resolution limit $\leq 3.0$ Å, but
	excluding above.
	SAD phasing: built residues/estimated residues $\geq$ 70%, but excluding above.
	Data reduction: R meas of inner resolution shell $\leq 15\%$ , high-resolution limit $\leq 3.5$ Å, but excluding above.
	SAD phasing: built residues/estimated residues $\geq$ 60%, but excluding above.
	All other successful jobs
	Job finished, but there was error happened when parsing the log files.
8	Job failed
X	Job is in the queue and waiting for executing
2	Job is executing
8	Unknown