Acta Crystallographica Section F

Volume 70 (2014)

Supporting information for article:

Structure of glutathione *S*-transferase-1 from the major human hookworm parasite *Necator americanus* (*Na*-GST-1) in complex with glutathione

Oluwatoyin A. Asojo and Christopher Ceccarelli

S1. Soaking studies

About 30 crystals were soaked with varying concentrations of a selected compound. Soaking with hematin resulted in complete loss of diffraction and clouding of the crystals. Almost all soaking experiments at higher than 100 μM final compound concentration resulted in the disintegration of the crystals, which cracked as soon as they came in contact with the soaking solution. Crystals soaked with 100 μM zinc protoporphyrin or protoporphyrin IX disodium salt were slightly yellow, and were stable at room temperature for at least 2 months.

S2. Data collection for zinc protoporphyrin soaked crystal

The data were processed and collected using CrysAlis^{Pro} and statistics are shown in Table S1. Data were collected using an Agilent Xcalibur PX Ultra four-circle kappa platform with a 165 mm diagonal Onyx CCD detector and a high-brilliance sealed-tube Enhance Ultra (Cu) X-ray source operating at 50 kV and 40 mA. Data collection was performed using the CrysAlis^{Pro} software (Agilent Technologies, 2012) to automatically select the best orientation of the crystal, while keeping the longest axis fixed to maximize spot separation. The process of keeping the longest axis fixed simply involved setting a single parameter in the CrysAlis^{Pro} software. The crystal to detector distance selected was 100 mm and exposure times of 45 seconds for 0.2° frame width and at -180^{o}C. Diffraction images showing the crystal quality are illustrated in Figure S1.

S3. The structure of *Na*-GST-1 soaked with zinc protoporphyrin

Native *Na*-GST-1 typically crystallizes in the primitive orthorhombic spacegroup $P2_12_12_1$ with two homodimers per ASU. While the data from the primitive crystal form can be collected and indexed on other X-ray instruments, the process was laborious and required visually attempting to mount the crystals to avoid the thinnest face as well as increasing the detector distance to 200 mm on home sources that were not equipped with four-circle kappa platform. Despite being a laborious process data collection process, the structure of apo *Na*-GST-1 was solved and refined to 2.4 Å (Asojo *et al.*, 2007). Soaking

these crystals with zinc protoporphyrin changes the space group to $C222_1$ with one homodimer per ASU. While the space group changes upon soaking, the cell dimensions are about the same, 48 x 80 x 200 Å.

The C-centered orthorhombic crystals were obtained by soaking primitive crystals with different compounds. We were unable to integrate or process any of the data of C-centered crystals collected on our local home source X-ray instruments, that were not equipped with a four-circle Kappa goniometer, likely because we could not successfully mount them visually to avoid the thinnest face. However, by keeping a single parameter fixed in the CrysAlis^{Pro} software we were able to implement a simple strategy for collecting data (Figure S

2). The strategy of keeping the longest axis fixed can be adopted with other systems that have a four-circle Kappa goniometer, and similar data set could have been collected at many synchrotron beamline stations. However the ease of implementation is dependent on the data collection software.

The long axis was resolved for the data set collected using only a 100 mm detector distance by collecting narrow frames (0.2° frame width) and by making use of a key feature of the automated data-collection strategy routine of CrysAlis^{Pro}. The strategy was constrained to position the four-circle Kappa goniometer such that the closely spaced c* axis was in the detector plane in as many of the diffraction images as possible. One or two additional orientations were required in order to collect a complete data set. The intensity statistics of are shown in Table S1. This data collection strategy ensured spot resolution along the c* axis, as is evident from Figure S1.

The C-centered orthorhombic structure was solved by molecular replacement with PHASER (McCoy *et al.*, 2005, Storoni *et al.*, 2004), using a monomer of *Na*-GST-1, pdb code 20N7 (Asojo *et al.*, 2007) stripped of waters and all ligands as the search model. Molecular replacement was followed by iterative cycles of manual model building with the program COOT (Emsley *et al.*, 2010) and structure refinement with REFMAC5 (Murshudov *et al.*, 2011) within the CCP4 package (Winn *et al.*, 2011). The resulting structure had no bound ligands and was refined to 2.6Å resolution. Coordinates and structure factors have been deposited in the rcsb protein databank under the entry code 4OFT. Structural data are shown in Table S2.

The structure is a prototypical GST dimer. The rmsd deviation of alignment of all main chain and side chain atoms of the C-centered orthorhombic dimer with the dimers of the search model was approximately 0.4 Å. The regions with the greatest difference in both structures are in proximity to the glutathione binding site (G-site) and cavity, which is suggestive of some minor conformational changes upon soaking with the ligand. In the unsoaked *Na*-GST-1 structure, the G-site is occluded by Gln 50 and the ability to bind glutathione is prevented, however upon soaking the conformation of the Gln 50 is changed as are the key residues in the active site that are involved in glutathione binding. Despite these changes, active site of the soaked crystals does not have electron density suggestive of any ligand being bound (Figure S3).

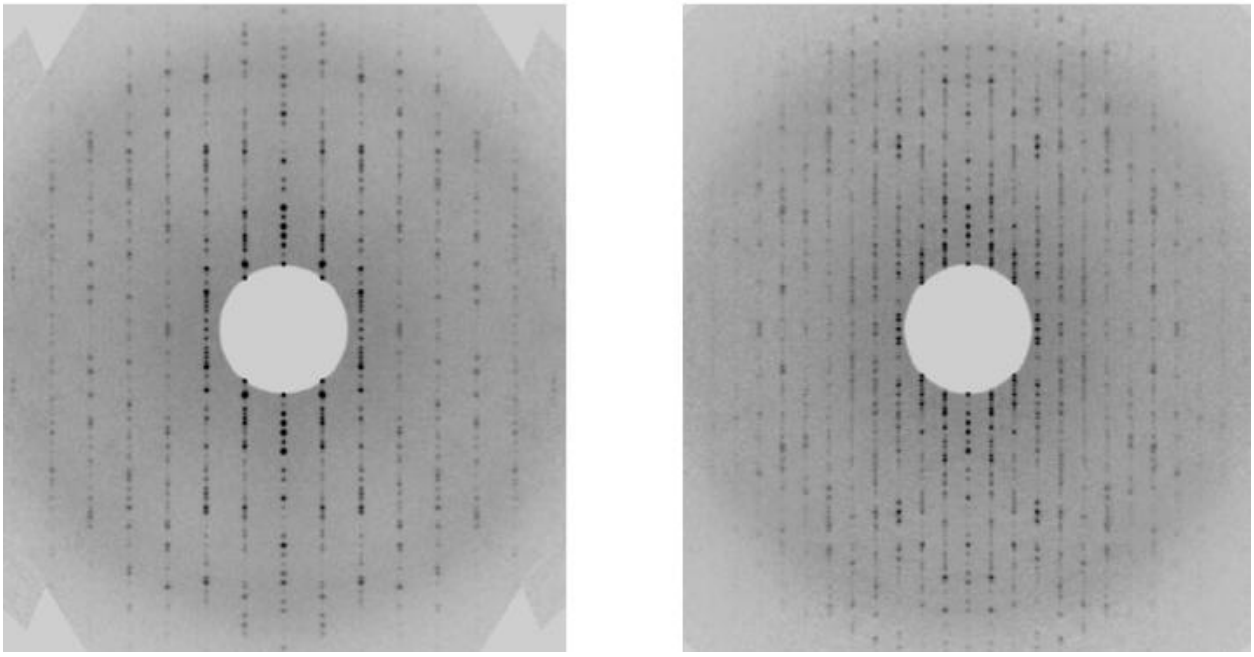

Figure S1 Spot separation is evident in reconstructed (A) h0l and (B) 0kl precession images shown along the c* vertical axis.

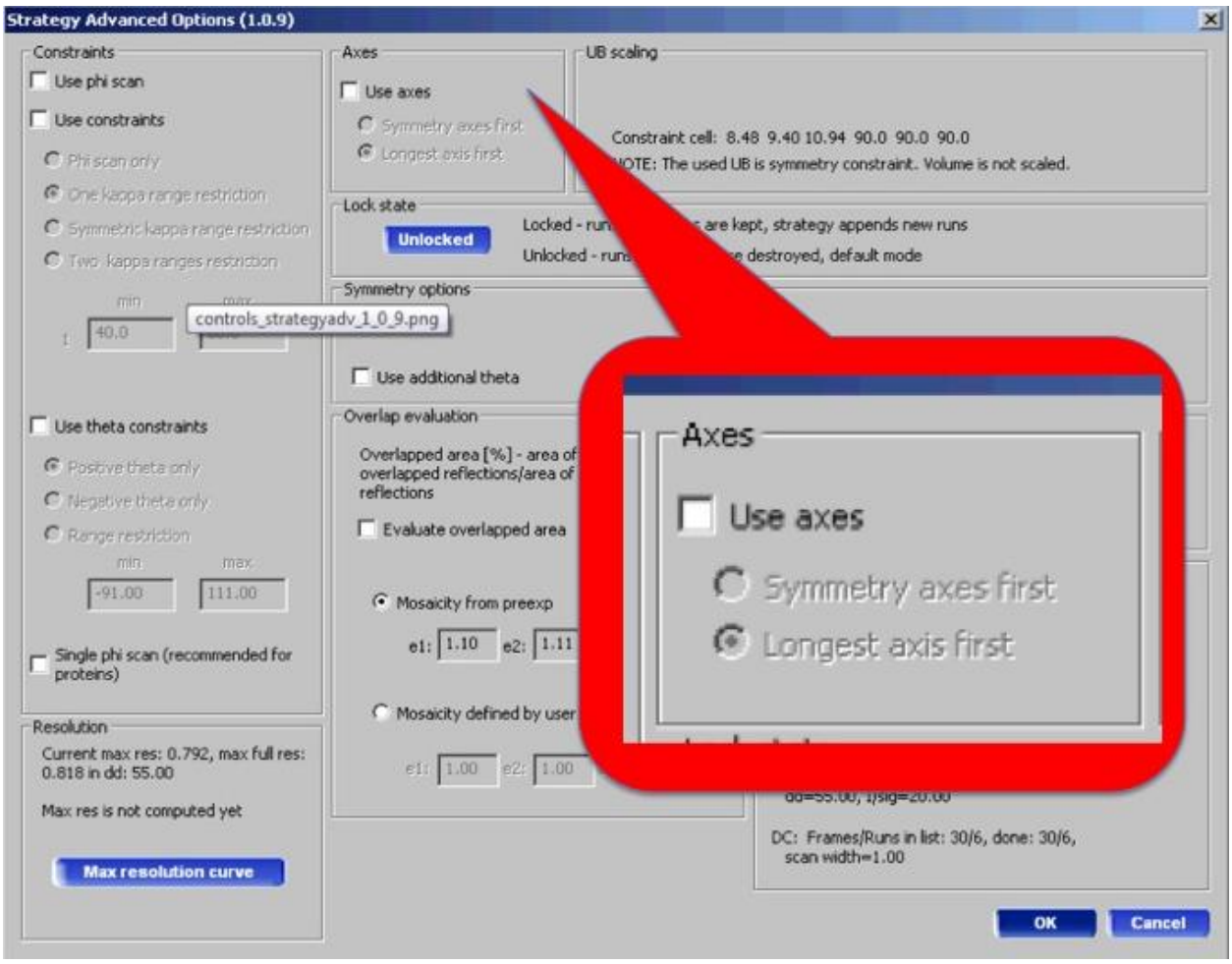

Figure S2 Screen grab showing the CrysAlis^{Pro} software option for fixing keeping longest axis fixed.

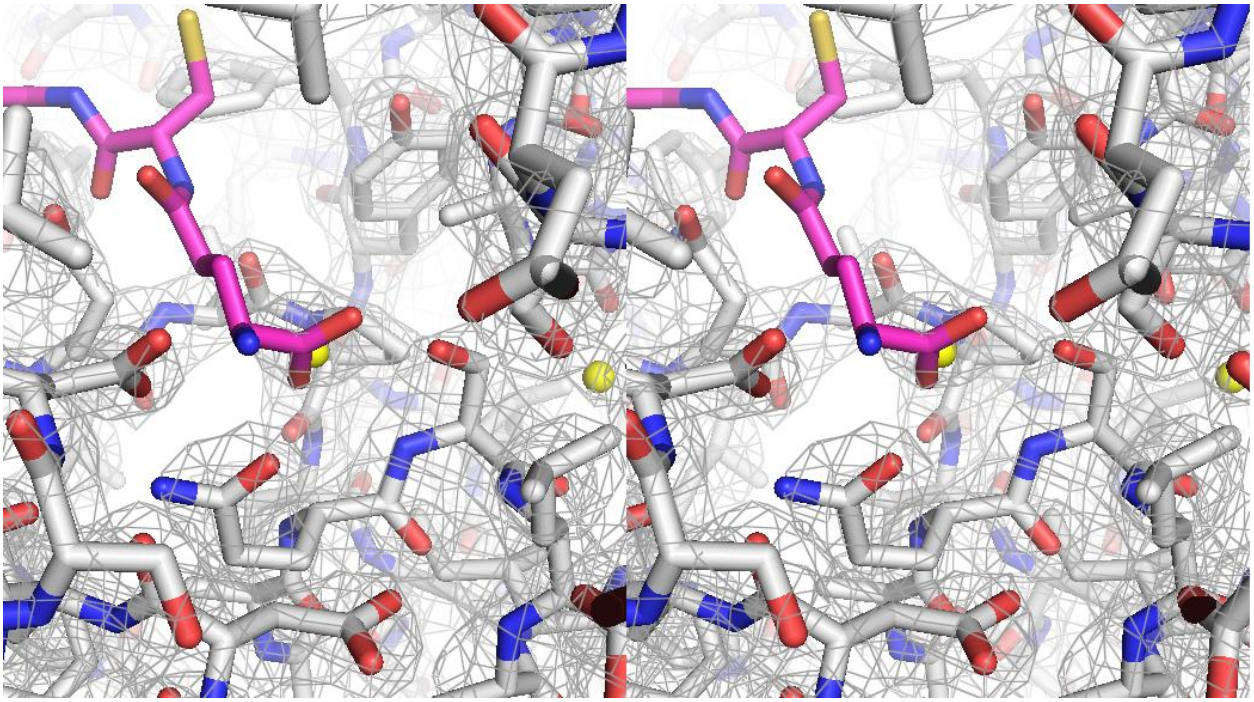

Figure S3 Stereo figure of the G-site of orthorhombic $C222_1$ (4OFT) reveals the absence of electron density for GSH. The final REFMAC5-generated $2F_o$ - F_c maps for the refined model is shown in gray and contoured at 1.5 σ, with the final refined model shown in gray sticks. The expected position of GSH is shown in magenta stick.

Table S1 Orthorhombic form intensity statistics

	Reflections								
Resolution (Å)	measured	theory	unique	%Complete	redundancy	<I >	<I/**σ**(I) >	R(int)	R(**σ**)
∞-5.75	10895	1284	1262	98.3	8.6	180.19	44.92	0.047	0.017
5.74-4.52	11881	1262	1262	100.0	9.4	102.24	25.85	0.070	0.033
4.52-3.93	11972	1262	1262	100.0	9.5	100.98	22.88	0.082	0.039
3.93-3.56	12061	1262	1262	100.0	9.6	78.58	17.27	0.107	0.052
3.56-3.30	12035	1262	1262	100.0	9.5	54.49	12.00	0.150	0.073
3.30-3.09	12190	1262	1262	100.0	9.7	31.04	7.71	0.247	0.120
3.09-2.94	12090	1262	1262	100.0	9.6	20.90	5.61	0.341	0.169
2.94-2.81	11979	1262	1262	100.0	9.5	14.39	4.02	0.460	0.230
2.81-2.69	9920	1262	1262	100.0	7.9	10.42	2.86	0.568	0.344
2.69-2.60	7400	1271	1265	99.5	5.8	8.39	2.15	0.639	0.461
∞-2.60	112423	12651	12623	99.8	8.9	61.90	14.92	0.127	0.061

$R(\boldsymbol{\sigma}) = [\Sigma\sigma(F_o^2)] / [\Sigma F_o^2]$

$R(\text{int}) = [\Sigma\, |F_o^2 - F_o^2\,(\text{mean})|\,] / [\Sigma F_o^2]$

Table S2 Crystallographic data collection and refinement statistics

Values in parentheses are for the highest resolution shell.

Data	*Orthorhombic C222$_1$*
RCSB code	4OFT
Unit cell	a=48.63 Å, b=80.89 Å, c=200.53 Å
Resolution (Å)	24.1-2.6 (2.7-2.6)
R_{merge} (%)[a]	6.1 (46.1)
Number measured reflections	112423 (7400)
Number unique reflections	12623 (1265)
Completeness (%)	99.8 (99.5)
Redundancy	8.9 (5.8)
I /σ(I)	14.9 (2.2)
Refinement Statistics	
R_{factor} (%)[b]	22.7 (26.1)
R_{free} (%)[c]	26.9 (29.9)
Number of atoms in model	
Protein	3348
Water	46
GSH	0
Mean B-factor (Å^2)	28.08
Rms deviation from ideal	
Bond length (Å)	0.013
Bond angles (^{o})	1.634
Chiral (Å^3)	0.109
Ramanchandran (% residues)	

Favored regions (%)	96.57
Allowed regions (%)	2.7
Outlier regions (%)	0.74

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - [I(hkl)]| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ and $[I(hkl)]$ are the intensity of measurement of I and the mean intensity of the reflection with indices hkl, respectively.

[b] $R_{cryst} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$ where F_o are observed and F_c are calculated structure factors amplitudes.

[c] R_{free} set uses 5% of randomly chosen reflections.

Asojo, O. A., Homma, K., Sedlacek, M., Ngamelue, M., Goud, G. N., Zhan, B., Deumic, V., Asojo, O. & Hotez, P. J. (2007). *BMC structural biology* **7**, 42.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta crystallographica. Section D, Biological crystallography* **61**, 458-464.

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta crystallographica. Section D, Biological crystallography* **67**, 355-367.

Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta crystallographica. Section D, Biological crystallography* **60**, 432-438.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242.