



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 09:08 am BST

PDB ID : 3QMZ  
Title : Crystal structure of the cytoplasmic dynein heavy chain motor domain  
Authors : Cho, C.; Carter, A.P.; Jin, L.; Vale, R.D.  
Deposited on : 2011-02-07  
Resolution : 6.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

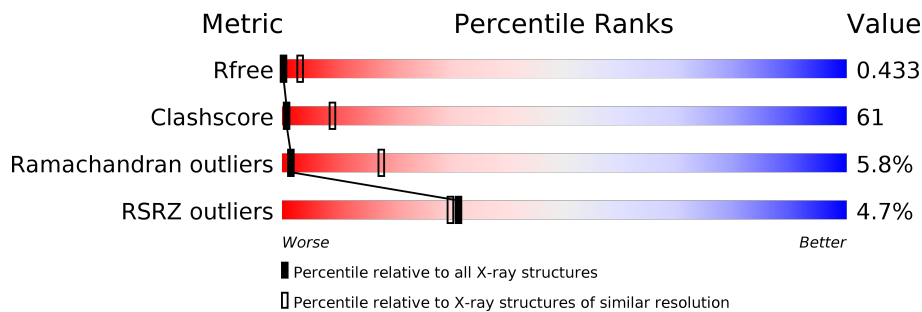
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 6.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1000 (8.00-3.88)
Clashscore	141614	1049 (8.00-3.90)
Ramachandran outliers	138981	1016 (8.00-3.86)
RSRZ outliers	127900	1015 (8.20-3.78)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	2486	
1	B	2486	
2	S	219	
2	T	219	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 23302 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Cytoplasmic dynein heavy chain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	2136	10585	6313	2136	2136	0	0	0
1	B	2136	10586	6314	2136	2136	0	0	0

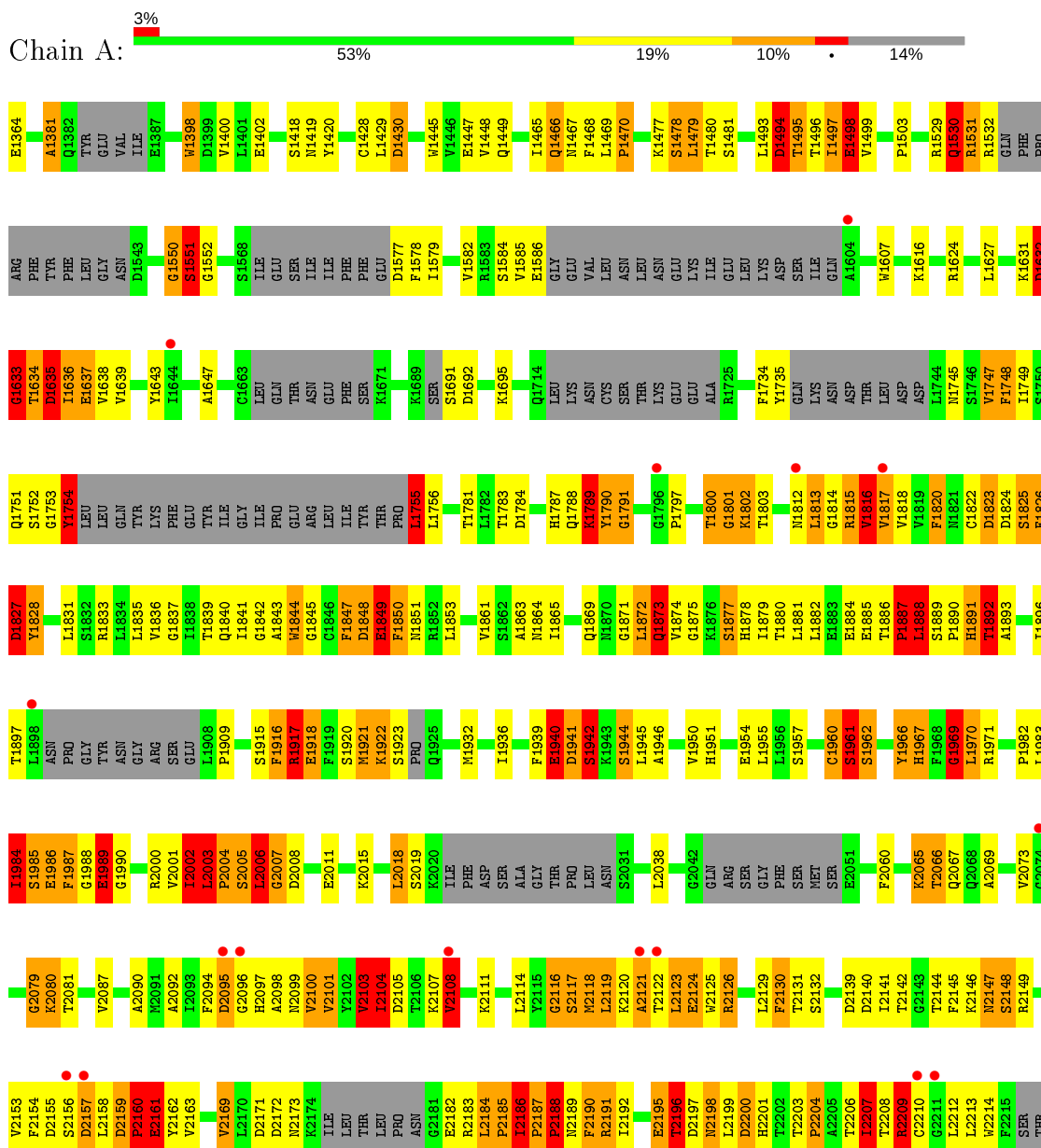
- Molecule 2 is a protein called Glutathione-S-transferase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	T	216	1066	634	216	216	0	0	0
2	S	216	1065	633	216	216	0	0	0

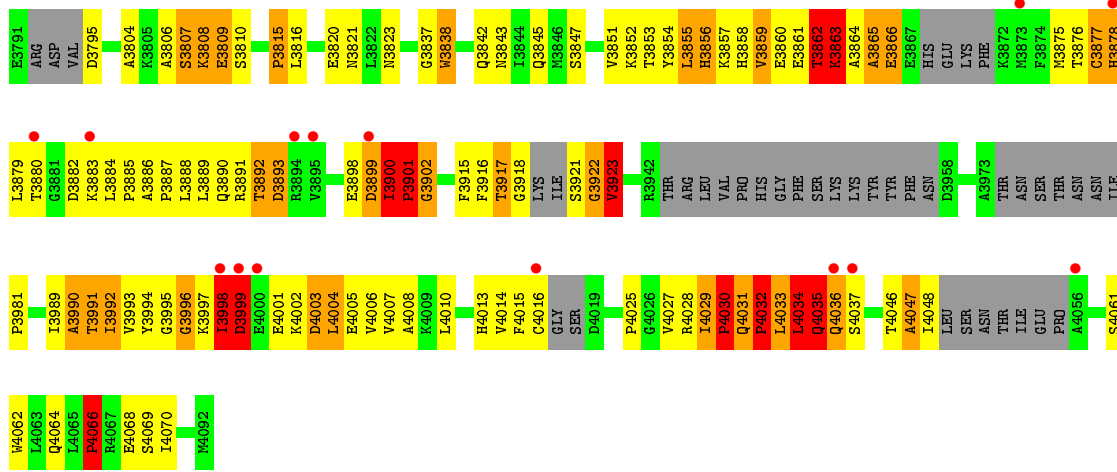
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

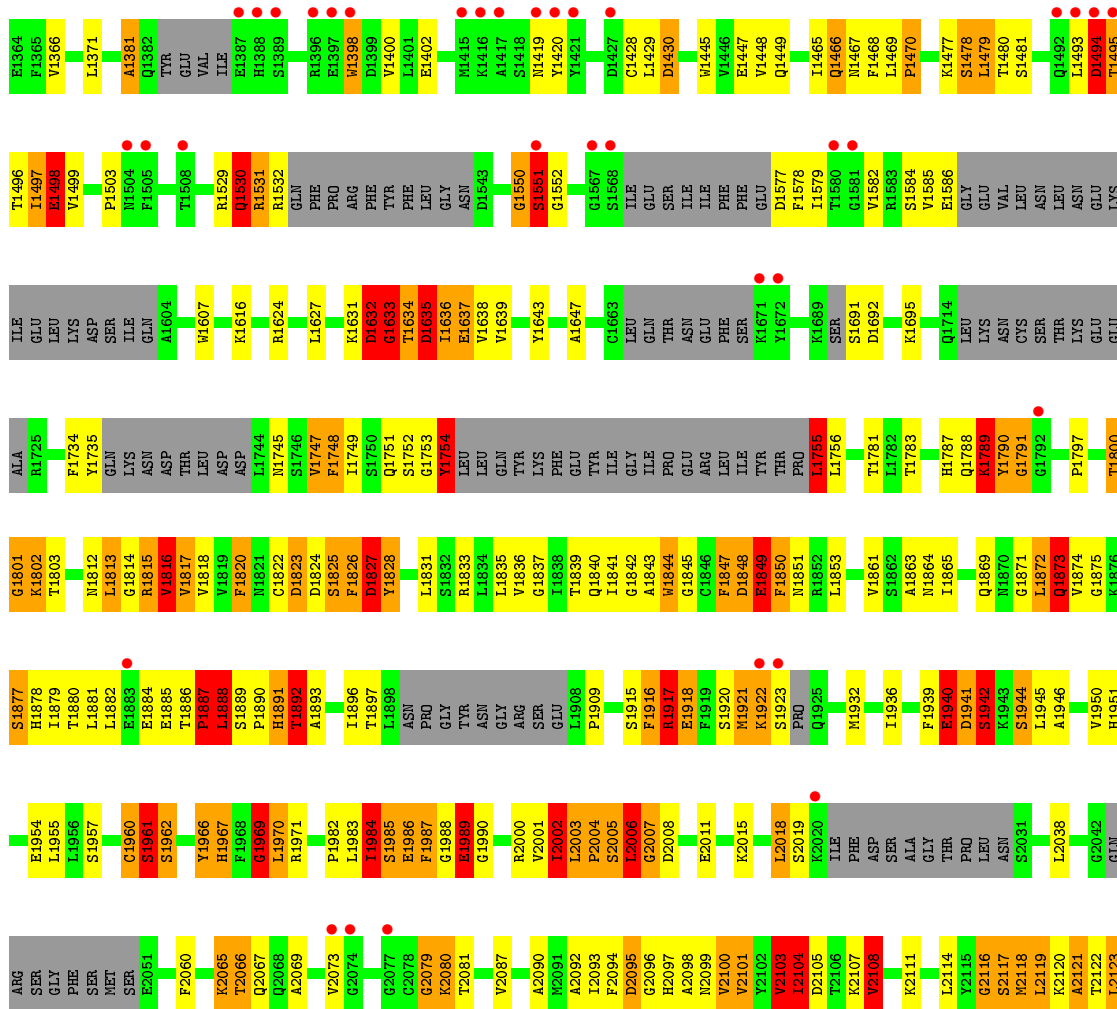
- Molecule 1: Cytoplasmic dynein heavy chain



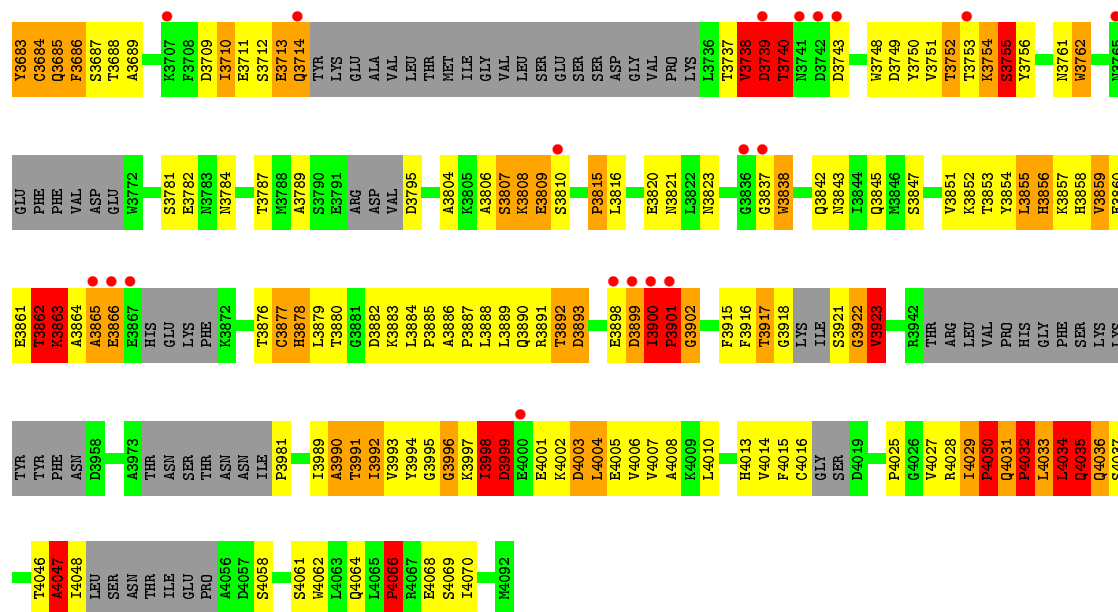




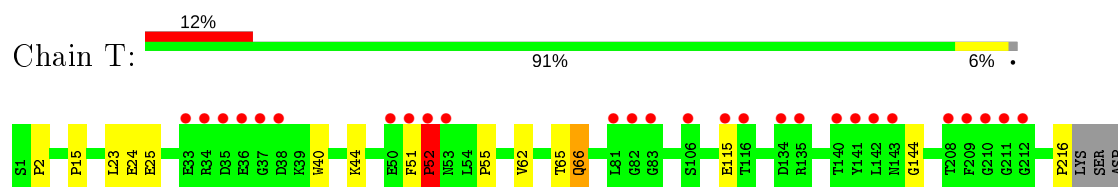
• Molecule 1: Cytoplasmic dynein heavy chain



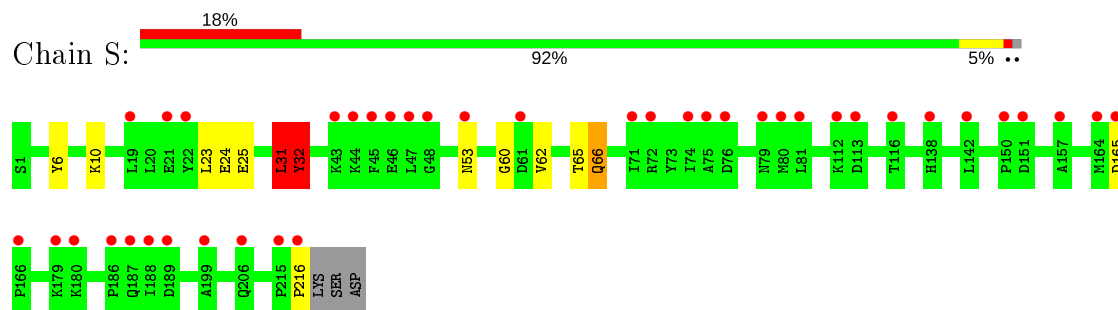
ASN	Q3453	I3362	LEU	E2872	I2708	L2616	G2533	F2371	A2291	E2195	E2124
LYS	D3454	I3363	ILE	L2873	R2709	F2617	A2534	C2372	A2295	E2196	W2125
GLU	G3455	I3364	HIS	Y2874	I2710	S2618	C2535	E2374	I2295	I2196	R2126
S2524	E3456	R3365	PHE	Z2875	D2875	P2619	N2536	E2375	W2301	D2197	L2123
	F3457	R3366	ASP	W2876	C2724	R2627	P2537	I2376	P2302	N2198	F2154
R3545	F3458	I3367	ARG	W2876	D2725	R2627	P2538	P2376	P2303	L2199	D2155
	D3459	I3370	ASN	I2881	D2726	R2627	P2539	S2377	Q2303	D2200	S2156
I3549	P3460	I3371	PHE	A2882	E2727	A2632	D2540	V2378	M2304	H2201	L2158
	L3461	T3372	TYR	K2883	LEU	I2633	R2541	S2379	L2305	T2202	L2163
R3565	L3462	L3373	GLN	N2634	GLU	N2634	R2542	L2380	D2306	P2204	E2168
E3568	S3463	L3374	LYS	P2637	VAL	P2637	R2543	E2381	D2307	T2203	M2169
E3569	R3464	P2731	MET	R2638	P2731	R2638	I2544	A2382	K2308	A2205	L2141
L3570	L3465	W2732	VAL	Q2639	W2732	Q2639	P2545	H2383	S2309	T2206	T2142
	I3466	I2733	GLY	T2640	I2733	T2640	R2546	E2384	L2310	I2207	F2145
F3470	W3379	I2734	GLY	L2641	I2734	L2641	S2547	W2385	K2311	T2208	K2146
ASN	R2898	Q2751	VAL	L2644	Q2751	L2644	R2551	M2386	R2209	R2209	N2147
HIS	S2899	K2757	PRO	E2651	R2757	E2651	R2552	P2388	G2211	G2211	S2148
ALA	P2906	L2758	ARG	A2652	L2758	A2652	H2553	D2389	L2212	L2212	R2149
GLY	A2907	I2759	SER	R2653	I2759	R2653	A2554	V2390	L2213	L2213	V2153
N3475	F2909	G2760	P2989	W2654	G2760	W2654	I2556	P2391	F2215	F2215	F2154
R3476	L2910	R2911	Q3014	R2654	R2910	R2654	LEU	P2393	SER	SER	D2155
F3477	ASN	R2911	Q3014	E2656	ASN	E2656	LEU	T2394	THR	THR	S2156
	ASP	R2911	Q3014	R2656	R2911	R2656	LEU	L2395	ASP	ASP	D2157
T3478	ASP	W2774	L3269	A2657	W2774	A2657	LEU	D2386	VAL	VAL	L2158
E3480	Y3389	L2776	LYS	ASP	L2776	ASP	GLY	T2397	S2334	S2334	D2159
I3481	N3399	G2777	VAL	ARG	G2777	ARG	ASP	Q2335	Q2335	Q2335	P2160
G3482	S3400	I2915	ASN	LEU	I2915	LEU	ASP	P2398	SER	SER	E2161
D3483	Q3401	W2916	GLU	VAL	W2916	VAL	ASP	K2399	R2336	R2336	E2161
H3484	ASP	W2917	LEU	VAL	W2917	VAL	ASP	L2407	A2337	A2337	Y2162
E3485	ALA	W2917	LEU	GLY	K2785	GLY	ASP	L2407	Q2340	Q2340	E2163
E3486	ALA	I2786	LYS	ASP	I2786	ASP	TRP	S2410	T2341	T2341	A2165
F3486	F3404	E2789	THR	ASP	E2789	E2789	ASP	K2411	N2339	N2339	M2166
ASP	F3405	T2922	THR	ASP	T2922	T2922	ASP	N2343	K2240	K2240	M2166
VAL	F3406	T2922	LEU	ASP	T2922	T2922	ASP	T2344	L2241	L2241	M2167
SER	L3407	T2922	PRO	ASP	R2812	R2812	ASP	Y2345	S2242	S2242	S2168
GLY	L3408	T2922	LYS	ASP	R2812	R2812	ASP	F2346	M2243	M2243	V2163
ASP	D3409	T2922	ALA	ASP	S2820	S2820	ASP	M2428	G2347	G2347	E2164
ASP	D3409	T2922	ALA	ASP	S2820	S2820	ASP	N2429	H2348	H2348	D2171
PHE	P3440	T2922	PRO	ASP	S2820	S2820	ASP	D2349	D2349	D2349	M2173
	S3411	T2922	PRO	ASP	S2820	S2820	ASP	E2352	E2352	E2352	K2174
L3494	S3412	T2922	GLU	ASP	S2820	S2820	ASP	D2255	D2255	D2255	ILE
F3495	H3413	T2922	LYS	ASP	S2820	S2820	ASP	N2432	N2432	N2432	LEU
I3496	H3413	T2922	LYS	ASP	S2820	S2820	ASP	N2433	N2433	N2433	THR
H3497	S3419	T2922	LYS	ASP	S2820	S2820	ASP	N2434	N2434	N2434	LEU
S3498	I3420	T2922	LYS	ASP	S2820	S2820	ASP	S2435	S2435	S2435	PRO
C3499	I3421	T2922	LYS	ASP	S2820	S2820	ASP	L2437	L2437	L2437	ASN
D3500	Y3422	T2922	GLU	ASP	S2820	S2820	ASP	V2441	V2441	V2441	G2181
P3501	Y3422	T2922	GLU	ASP	S2820	S2820	ASP	G2442	G2442	G2442	E2182
S3502	F3334	T2922	GLU	ASP	S2820	S2820	ASP	K2442	K2442	K2442	R2183
G3503	F3334	T2922	GLU	ASP	S2820	S2820	ASP	N2444	N2444	N2444	L2184
D3504	G3335	T2922	GLU	ASP	S2820	S2820	ASP	F2445	F2445	F2445	P2185
I3505	H3336	T2922	GLU	ASP	S2820	S2820	ASP	S2446	S2446	S2446	I2186
P3506	L3428	T2922	GLU	ASP	S2820	S2820	ASP	K2447	K2447	K2447	P2187
	L3429	T2922	GLU	ASP	S2820	S2820	ASP	E2452	E2452	E2452	P2188
R3514	S3430	T2922	GLU	ASP	S2820	S2820	ASP	L2284	L2284	L2284	N2189
L3515	F3431	T2922	GLU	ASP	S2820	S2820	ASP	Q2289	Q2289	Q2289	F2190
V3516	F3431	T2922	GLU	ASP	S2820	S2820	ASP	S2369	S2369	S2369	R2191
H3517	F3432	T2922	GLU	ASP	S2820	S2820	ASP	S2370	S2370	S2370	I2192
E3494	E3433	T2922	VAL	ASP	S2820	S2820	ASP				
E3495	E3434	T2922	VAL	ASP	S2820	S2820	ASP				
I3518	E3435	T2922	VAL	ASP	S2820	S2820	ASP				
F3519	E3435	T2922	VAL	ASP	S2820	S2820	ASP				
T3520	I3452	T2922	VAL	ASP	S2820	S2820	ASP				



- Molecule 2: Glutathione-S-transferase



- Molecule 2: Glutathione-S-transferase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.08Å 118.92Å 200.51Å 90.00° 90.27° 90.00°	Depositor
Resolution (Å)	50.00 – 6.00 50.13 – 6.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-6.00) 99.9 (50.13-6.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.38 (at 6.15Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.430 , 0.430 0.440 , 0.433	Depositor DCC
$R_{free}$ test set	1079 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	259.8	Xtrriage
Anisotropy	0.257	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 756.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.014 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.72	EDS
Total number of atoms	23302	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	222.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality i

### 5.1 Standard geometry i

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	1.63	201/10536 (1.9%)	2.24	571/14615 (3.9%)
1	B	1.63	201/10537 (1.9%)	2.24	572/14617 (3.9%)
2	S	1.94	2/1064 (0.2%)	0.86	3/1479 (0.2%)
2	T	0.39	0/1065	0.78	6/1481 (0.4%)
All	All	1.61	404/23202 (1.7%)	2.15	1152/32192 (3.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	3	136
1	B	3	136
2	S	0	3
All	All	6	275

The worst 5 of 404 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	S	32	TYR	N-CA	49.28	2.44	1.46
2	S	32	TYR	CA-C	37.84	2.51	1.52
1	B	4047	ALA	C-N	-30.16	0.64	1.34
1	A	4047	ALA	C-N	-30.14	0.64	1.34
1	A	3426	THR	C-N	-25.27	0.76	1.34

The worst 5 of 1152 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1466	GLN	O-C-N	-31.44	72.39	122.70
1	A	1466	GLN	O-C-N	-31.43	72.42	122.70
1	B	2436	SER	O-C-N	-28.51	77.09	122.70
1	A	2436	SER	O-C-N	-28.49	77.12	122.70

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1466	GLN	CA-C-N	-28.26	55.02	117.20

5 of 6 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1888	LEU	CA
1	A	2375	ILE	CA
1	A	3641	PHE	CA
1	B	1888	LEU	CA
1	B	2375	ILE	CA

5 of 275 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1381	ALA	Mainchain
1	A	1430	ASP	Mainchain
1	A	1479	LEU	Mainchain
1	A	1494	ASP	Mainchain,Peptide
1	A	1530	GLN	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10585	0	4454	1006	3
1	B	10586	0	4455	994	3
2	S	1065	0	464	9	0
2	T	1066	0	465	22	0
All	All	23302	0	9838	2013	3

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 61.

The worst 5 of 2013 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1364:GLU:CA	2:T:216:PRO:HA	1.19	1.65
1:A:1932:MET:CB	1:A:1946:ALA:HB1	1.24	1.64
1:B:2060:PHE:CB	1:B:2087:VAL:CB	1.78	1.62
1:A:2060:PHE:CB	1:A:2087:VAL:CB	1.78	1.57
1:A:3666:ALA:HB2	1:A:3668:ARG:CB	1.32	1.57

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2245:GLU:N	1:B:2241:LEU:CB[1_554]	1.43	0.77
1:A:2245:GLU:CA	1:B:2241:LEU:CB[1_554]	1.81	0.39
1:A:2245:GLU:CB	1:B:2241:LEU:CB[1_554]	2.05	0.15

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	2039/2486 (82%)	1605 (79%)	307 (15%)	127 (6%)	1	16
1	B	2039/2486 (82%)	1605 (79%)	308 (15%)	126 (6%)	1	16
2	S	214/219 (98%)	194 (91%)	16 (8%)	4 (2%)	8	38
2	T	214/219 (98%)	190 (89%)	20 (9%)	4 (2%)	8	38
All	All	4506/5410 (83%)	3594 (80%)	651 (14%)	261 (6%)	1	17

5 of 261 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1470	PRO
1	A	1494	ASP
1	A	1498	GLU
1	A	1635	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1637	GLU

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 5.7 Other polymers [i](#)

There are no such residues in this entry.

### 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	109
1	A	109

The worst 5 of 218 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1756:LEU	C	1776:LEU	N	12.66
1	B	1756:LEU	C	1776:LEU	N	12.66
1	A	3028:VAL	C	3289:LEU	N	11.45
1	B	3028:VAL	C	3289:LEU	N	11.45
1	A	3981:PRO	C	3982:TRP	N	5.84

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	2136/2486 (85%)	-0.03	69 (3%) 47 41	183, 210, 384, 480	0
1	B	2136/2486 (85%)	0.07	84 (3%) 39 34	183, 210, 384, 480	0
2	S	216/219 (98%)	0.96	40 (18%) 1 3	209, 233, 253, 268	0
2	T	216/219 (98%)	0.66	27 (12%) 3 7	201, 219, 248, 273	0
All	All	4704/5410 (86%)	0.09	220 (4%) 31 30	183, 210, 376, 480	0

The worst 5 of 220 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	S	80	MET	7.1
1	B	2156	SER	6.0
1	B	1397	GLU	5.5
1	B	3899	ASP	5.3
2	T	52	PRO	5.2

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.